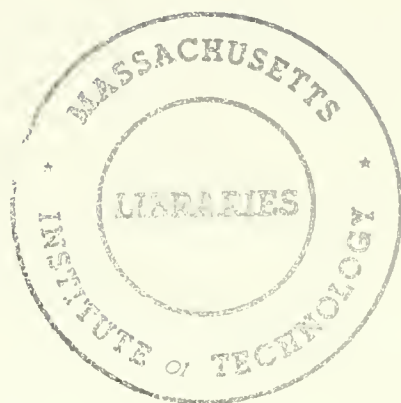


BASEMENT

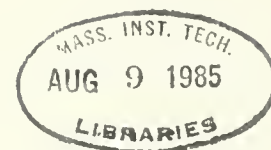






HD28  
.M414  
no. 1670-  
85

Dewey.



WORKING PAPER  
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

DIVIDEND BEHAVIOR FOR THE  
AGGREGATE STOCK MARKET

Terry A. Marsh  
Robert C. Merton

WP #1670-85

September 1983  
May 1985

MASSACHUSETTS  
INSTITUTE OF TECHNOLOGY  
50 MEMORIAL DRIVE  
CAMBRIDGE, MASSACHUSETTS 02139



DIVIDEND BEHAVIOR FOR THE  
AGGREGATE STOCK MARKET

Terry A. Marsh  
Robert C. Merton

WP #1670-85

September 1983  
May 1985



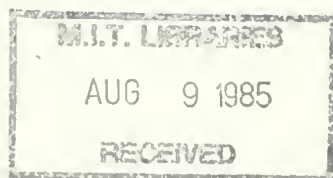


/ DIVIDEND BEHAVIOR FOR THE AGGREGATE STOCK MARKET\*<sup>†</sup> /

Terry A. Marsh  
Robert C. Merton

Sloan School of Management  
Massachusetts Institute of Technology  
50 Memorial Drive  
Cambridge, MA 02139

First Draft: September 1983  
Last Revised: May 1985



## 1. Introduction

In this paper, we develop a model of the dividend process for the aggregate stock market which fits the data rather well. Previous research has focused almost exclusively on dividend behavior at the micro level of the individual firm. Hence, to motivate the focus here on aggregate dividend behavior, we begin with a brief review of these earlier micro studies, this to be followed by a discussion which locates the place of our aggregate analysis within this body of research. In Sections 2-5, we derive and fit our econometric model of the dividend process. In Section 6, we compare the performance of the model with other models in the literature.

Although long a staple of financial management textbooks, corporate dividend policy remains a topic on which the field has failed to arrive at even a local sense of closure.<sup>1</sup> Fischer Black (1976) has aptly described this lack of closure as the "dividend puzzle." The pivotal point in this puzzle is the classical work of Miller and Modigliani (1961) which demonstrated the irrelevance of dividend policy for determining the firm's cost of capital.

Miller and Modigliani showed that when investors can create any payout pattern they want by selling and purchasing shares, the expected return required to induce them to hold these shares will be invariant to the way in which firms "package" gross dividend payments and new issues of stock (and/or other zero net present value transactions). Since neither the firm's expected future net cash flows nor its discount rate is affected by the choice of dividend policy per se, its current market value cannot be changed by a change in that policy. Thus dividend policy "does not matter." Although, under the MM proposition, there are no a priori reasons for firms to follow any

systematic dividend policy, there are also no penalties if they choose to do so.

Exceptions to the MM view are, of course, to be found in the literature. Gordon (1959;1962) and Lintner (1962) claimed that dividend policy does affect the firm's cost of capital, and provided some early evidence to support the view that a higher dividend payout reduces the cost of capital (i.e., investors prefer dividends). Others argue that personal and corporate taxes cause dividend policy to affect the firm's cost of capital, but in the direction that a higher payout raises the cost of capital (i.e., investors prefer capital gains), and Litzenberger and Ramaswamy (1979) and Poterba and Summers (1984) offer empirical support for this view. On the other hand, Black and Scholes (1974), Miller and Scholes (1978;1982), Hess (1983), and Eades, Hess and Kim (1984) present analysis and evidence suggesting that, as an empirical matter, tax effects per se do not appear to affect the cost of capital. Along different lines from these studies, Shefrin and Statman (1984) have used behavioral theories of individual choice to argue that investors will prefer cash dividends, even if they are tax disadvantaged. Although some of these analyses might provide reasons to believe that investors are not indifferent between cash dividends and capital gains, the empirical evidence to date is still inconclusive for rejecting the Miller and Modigliani proposition.

Even with their view of investor indifference for dividends, Miller and Modigliani (1961, p. 431) did point out that dividend policy could matter if dividend changes were used by firms to convey information not otherwise known to the market. Bhattacharya (1979), Miller and Rock (1982), and John and Williams (1984) use a signalling model approach to formalize this notion. Aharony and Swary (1980) provide some empirical evidence of the informational

content of dividends in their study of dividend announcement events for 149 NYSE industrials, in which they find that, on average, unexpected dividend and unexpected price changes are positively correlated around announcement dates. Asquith and Mullins (1983) find similar results for firms that initiate dividend payments for the first time. The evidence is, however, that the correlations, while statistically significant, are rather small, a conclusion also reached in empirical work by Watts (1973) and Gonedes (1978).

In summary, there are a number of conflicting theories of dividend behavior, and the empirical studies to date provide little compelling evidence for one over the others. The management of a firm is free to choose a dividend policy with virtually any time pattern it wants, subject only to the overall constraint that the present value of expected future distributions net of new stock offerings cannot exceed the present value of the firm's expected net cash flows generated by its investments. Indeed, except for certain debt indenture restrictions and accumulated earnings tax penalties, there do not appear to be any significant legal, accounting convention, or corporate tax factors to exert pressures on managers of publicly-traded and widely-held corporations to follow any particular dividend policy.

With so much controversy surrounding the various normative theories of the dividend process, it is perhaps not surprising that empirical researchers have relied heavily on positive theories of dividend behavior to specify their models. The prototype for these models is the Lintner model (1956) which is based on stylized facts first established by him in a classic set of interviews of managers about their dividend policies. From these interviews, it was readily apparent that dividend policies across firms were hardly uniform. Lintner did, however, identify some common characteristics: Namely,

managers tend to change dividends primarily in response to an unanticipated and nontransitory change in their firms' earnings, and they are guided by target payout ratios in making those changes. Using an econometric model based on these perceived patterns, Lintner found that he could explain a significant portion of annual dividend changes for a sample of companies over the period 1918-1941. Using similar types of models, subsequent empirical work by Fama and Babiak (1968), Petit (1972), and Watts (1973) supports Lintner's original findings.

With few exceptions (notably, Brittain (1966) and Shiller (1981a,b)), research on both normative and positive models of dividends has focused on the micro behavior of the individual firms. The relative lack of research on aggregate dividend behavior is perhaps not surprising since many of the more interesting issues surrounding dividend policy are likely to be firm specific. For example, clientele effects and indenture restrictions which could in principle affect an individual firm's dividend policy are likely to "wash out" in any aggregate dividend analysis. Similarly, issues involving the informational content of an individual firm's dividends are likely to be considerably less important for the stock market as a whole than for an individual firm. It is, indeed, difficult to see how one could identify meaningful announcement dates for aggregate dividends to perform event studies along the lines of (say) Gonedes (1978) and Aharony and Swary (1980).

If firms only changed their dividends to signal information and if the only information worth signalling is specific to the firm, then changes in aggregate dividends would be random and their magnitudes small. If, however, firms change their dividends for reasons other than signalling, then the very fact that aggregate dividend changes are unlikely to contain much signalling



information may make them an especially useful series for measuring the informational content of an individual firm's dividend announcements. To identify the signals or abnormal changes in a firm's dividends, it is, of course, necessary to have a model of its "normal" dividend behavior. Watts (1973), for example, used the Lintner model for this purpose. The Lintner model does not, however, take account of the cross-sectional dependencies among firms' dividend policies. It is reasonable to expect that in addition to its own economic circumstances, the firm would use the dividend behavior of other firms to calibrate its dividend policy--as, for example, observing industry practice in the selection of its target payout ratio. Moreover, these dependencies may be of considerable empirical significance in light of the already-documented strong correlations among different firms' contemporaneous stock price changes. Removal of the aggregate market component of a stock's return to obtain a better estimate of its abnormal price change is commonplace. Just so, use of an aggregate dividend model to remove the "systematic" component of an individual firm's dividend policy would appear to provide a better estimate of its abnormal dividend changes.<sup>2</sup>

If individual firms follow reasonably stable dividend policies over time, then the afore-mentioned cross-sectional dependencies will induce systematic behavior in the time series of aggregate dividends. It is, however, possible for aggregate dividends to exhibit stable and consistent time series properties even if no such stability were found for individual firms. To see this, consider the following polar case of a completely "demand-driven" model for dividends.

Suppose that firms have no need to use dividend policy to convey information to the market. Suppose further that some investors prefer at

least some portion of their total returns in the form of dividends. A fundamental assumption of portfolio theory is that investors are only concerned about the return characteristics (including, in general, the mix of dividends and capital gains) on their optimal portfolios. That is, if the menu of return characteristics available in the set of all optimal portfolios is fixed, then investors will be indifferent as to the way in which these portfolio characteristics are generated by the individual securities from which they are constructed.<sup>3</sup> Thus, investors only care about the dividend-capital gain mix at the portfolio level, and therefore, their demand for dividends is not "firm specific." Hence, in general, there will be many different allocations of dividend policies to individual firms that will support an equilibrium in the "dividend market."

As already noted, there are few institutional restrictions on a firm's choice of dividend policy. Hence, virtually all firms with accumulated earnings are potential suppliers of dividends. If, as assumed in this polar example, dividends are not used to convey information, then firms can change their dividend policies without creating "false signals." With no apparent barriers to entry (or exit), it seems reasonable to assume that the market for supplying dividends is competitive. As we all know, in a competitive industry, the equilibrium aggregate supply is determinate, but the number of firms in the industry and the quantity each produces are largely indeterminate. Just so, equilibrium aggregate dividends may be determinate, but which firms service this demand and the quantity each chooses to supply may not.

In such an environment, the dynamic behavior of aggregate dividends can be quite systematic even if the individual dividend policies of firms exhibit



relatively little intertemporal stability. Although consistent with such micro instability, this dividend story certainly does not require that the time series of individual firms' dividends have no predictable properties. Indeed, if the story were slightly embellished to include transactions costs for either issuers or investors, then even this extreme demand-driven model would predict systematic micro dividend behavior. With issuing costs, for example, firms in mature or declining industries with large cash flows relative to their new investment needs, would be marginally lower-cost producers of dividends than firms in growth industries.<sup>4</sup> Since the position of a firm or an industry within its "life cycle" is hardly random from year to year, both individual firm and industry payout patterns are likely to exhibit serial dependencies.

By now, the reader has surely recognized this dividend story as essentially a (incomplete) version of the Miller (1977) theory of corporate debt applied to dividends. Like the Miller theory with respect to the observed debt policies of individual firms, this story is consistent with the empirical fact that firms pursue widely different dividend policies, including the one of paying no dividends at all. As with the Miller theory, the strict version of our story assumes that firms do not use financing policy as an important means of conveying otherwise unknown information to the market. However, weakening of this assumption need not affect the behavior of aggregate dividends if either these signalling components of individual firms' dividend changes are sufficiently orthogonal to one another to "cancel out" in the aggregate or there are a sufficient number of firms with "passive" (i.e., nonsignalling) dividend policies so as to offset the effect on aggregate dividends of those firms which actively use dividend policy to transmit information.

As is readily apparent from our discussion, a particular stable pattern of aggregate dividend behavior will in general be consistent with more than one normative theory of micro dividend behavior. Lack of uniqueness does not, however, imply no restrictions. If, for example, aggregate dividends exhibit systematic behavior, then firm-specific theories such as signalling cannot alone explain the reasons why firms change their dividends. On the other hand, clientele models in which investors' transactions costs are nontrivial and reasonably stable over time will generally be consistent with the restrictions implied by the systematic behavior of aggregate dividends. Of course, it might seem that not only have transactions costs not been stable, but that they are relatively insignificant today. Yet such costs were once important, and together with the limited communications, limited data transmission mechanisms, and the low or nonexistent income tax rates of yesteryear, they may well have caused dividends to be an historically important component of asset returns. Although these inducements for dividend payments have themselves changed substantially, it is neither irrational nor rational for dividend payouts to have remained in existence so long as they have not become disadvantaged by recent developments. That is, those dividend payout patterns observed today could be "nonrationally" explained by factors which are now economically unimportant.<sup>5</sup> Moreover, historically-determined payout ratios should have long-run steady-state distributions unless events having a nonneutral impact on the aggregate quantity of dividends demanded cannot be absorbed by long-run supply adjustments, a possibility which seems unlikely.

Given the potential role of precedent in explaining dividends, and in the absence of an alternative well-formulated normative macro theory of dividends, we turn in the next section to the descriptive facts established in the Lintner interviews to motivate the specification of our econometric model.<sup>6</sup>

While sharing Lintner's stylized facts in common with the previously-cited empirical studies of micro dividend behavior, our model, in addition to being applied at the aggregate level, differs significantly from these earlier studies because it assumes that economic earnings, instead of accounting earnings, are the primary determinant of dividends. The analysis in Section 6 compares the performance of our model with one which uses accounting earnings. Because the Brittain (1966) study of aggregate dividends relies upon the relations between dividends and accounting variables, our analysis sheds light on his findings as well. In this same section, we also compare the relative performance of our model with the univariate trend-autoregressive model which is associated with the Shiller (1981a,b) model of aggregate dividends.

## 2. Model of Aggregate Dividend Dynamics

Lintner found considerable heterogeneity among firm dividend policies in his interviews of corporate managers. However, he also found some characteristics to be common to many of these firms' dividend policies. These stylized facts are summarized as follows: (i) Managers believe that firms should have some long-term target payout ratio; (ii) In setting dividends, they focus on the change in existing payouts and not on the level; (iii) A major unanticipated and nontransitory change in earnings would be an important reason to change dividends; (iv) Most managers try to avoid making changes in dividends which stand a good chance of having to be reversed within the near future.

Most textbook discussions seem to agree with the interpretation of these stylized facts to the effect that it is changes in some measure of long-run sustainable or "permanent" earnings, rather than current earnings, which drive dividend decisions. That is, a change in current earnings flow which is viewed by management as essentially transitory would not be likely to give rise to a noticeable change in dividends. Unfortunately, except for the special case of a firm whose future earnings are certain and generated without further net new investment, the textbooks are not specific in defining permanent earnings. Our interpretation (which is consistent with this special case) defines the permanent earnings per share of a firm at time  $t$  as equal to the expectation as of time  $t$  of that level of uniform payments which could be made by the firm to a single share in perpetuity. For an all-equity financed firm, permanent earnings are determined as follows: Let  $\Pi(s)$  denote the real after-tax cash flow from the physical and financial assets of the firm at time  $s$  and  $I(s)$  denote the real net new investment by the firm at time  $s$ .  $I(s) = [\text{gross new physical investment} + \text{purchases of financial}$

assets - sales of physical and financial assets]. If  $\alpha$  denotes the firm's real cost of capital, then the discounted value of the expected cash flows available for distribution to each share outstanding at time  $t$  is given by:

$$V(t) \equiv \epsilon_t \left\{ \int_t^{\infty} [\Pi(s) - I(s)] e^{-\alpha(s-t)} ds \right\} / N(t) \quad (1)$$

where  $\epsilon_t$  denotes the expectation operator, conditional on information available as of time  $t$  and  $N(t)$  denotes the number of shares outstanding.  $V(t)$  is sometimes called the "intrinsic value" (per share) of the firm, and permanent earnings per share are determined by creating a perpetual annuity from this intrinsic value.<sup>7</sup> That is, if  $E(t)$  denotes permanent earnings per share of the firm at time  $t$ , then:

$$E(t) = \alpha V(t) \quad (2)$$

Since corporate managers set dividends for their firms, it is their assessments of permanent earnings which are relevant for the evolution of aggregate dividends. For this purpose, we denote managers' determination of permanent earnings by  $E^m(t) = \alpha V^m(t)$  where  $V^m(t)$  is given by (1) with the expectation operator  $\epsilon_t = \epsilon_t^m$  based on the probability distribution for future  $\Pi(s)$  and  $I(s)$  generated by the managers' information sets as of time  $t$ .

Although Lintner's stylized facts suggest that dividend changes are related to permanent earnings changes, the interview data on which they are based contain little information about the detailed functional form of that relation.<sup>8</sup> In the absence of a specific structural model of that relation, we posit that logarithmic dividends can be expressed as the sum of a rational distributed lag of logarithmic permanent earnings, a drift term which is conditional on information known at time  $t$ ,  $a(t)$ , and a disturbance term  $\eta(t)$ . That is, we represent the aggregate dividend process as:



$$(1-\phi_1 L) \log[D(t)] = a(t) + (\lambda - \theta_1 L) \log[E^m(t-1)] + \eta(t) . \quad (3)$$

where  $D(t)$  is the integral  $\int_{t-1}^t D(s)ds$  of aggregate dividends paid per share of the market portfolio over the interval from time  $t-1$  to time  $t$ ;  $E^m(t)$  is permanent earnings as defined in (2), per share of the market portfolio at time  $t$ ; and the roots of the first order polynomials in the lag operator  $L$  are outside the unit circle.

As specified, (3) is consistent with the "short run" dividend dynamics of the Lintner model. It does not, however, capture his stylized fact (i) that firms typically set a long-run target for the dividend payout ratio. In line with the discussion concerning the steady-state properties of long-run equilibrium dividend payout, we take account of this long-run objective in our model by requiring that dividend payouts converge to a constant target ratio, i.e., as  $t \rightarrow \infty$  (and in the absence of any disturbances),

$$\lim_{t \rightarrow \infty} \log \left[ \frac{D(t)}{E^m(t-1)} \right] = \beta . \quad (4)$$

This special assumption that the long-run target be literally constant is more stringent than necessary.<sup>9</sup> Moreover, this assumption does not imply that dividends and permanent earnings follow (trend) stationary processes. As shown in Appendix A, the rate of payout  $D(t)/E^m(t)$  can tend to a steady state even though  $E^m(t)$  and  $D(t)$  may not.

It is shown in Appendix B that when the long-run steady-state (4) is imposed on the short-run dynamics (3), the latter can be rewritten as:

$$\begin{aligned} \log[D(t+1)] - \log[D(t)] = & g(t) + \lambda \{ \log[E^m(t)] - \log[E^m(t-1)] - m(t-1) \} \\ & + \gamma \{ \beta - (\log[D(t)] - \log[E^m(t-1)]) \} + \eta(t+1) . \end{aligned} \quad (5)$$

where  $m(t-1)$  is the time  $t-1$  expectation of the logarithmic change in

permanent earnings,  $\{\log[E^m(t)] - \log[E^m(t-1)]\}$ ; and  $g(t)$  is the logarithmic change in dividends,  $\{\log[D(t)] - \log[D(t-1)]\}$ , which would be expected given that the time  $t-1$  logarithmic payout ratio  $\log[D(t)/E^m(t-1)]$  is equal to its long-run target  $\beta$  and the unexpected change in logarithmic permanent earnings,  $\{\log[E^m(t)] - \log[E^m(t-1)] - m(t-1)\}$ , is zero.

The model described by (5) takes the form of the well-known "error correction" model which has been studied and applied by Sargan (1964), Davidson, Hendry, Srba, Yeo (1978), Nickell (1980), Salmon (1982), and Hendry and Ericsson (1984). It obviously satisfies the condition for a long-run steady state distribution for  $D/E^m$  because if  $\eta(t+1) = 0$ , then  $\Delta \log[E^m(t)] = m(t-1)$  implies  $\{\log[D(t)] - \log[E^m(t-1)]\} = \beta$  as required. Given the specification (4) of the long-run equilibrium, the model's potential for describing the short-run dynamics of aggregate dividends depends upon the appropriateness of the rational distributed lag in (3). Since the error correction model is applicable for a wide range of stochastic processes governing  $E^m(t)$ , including the random walk, (cf. Appendix A and Nickell (1980)), the major assumptions imbedded in (3) are those of symmetry in the responsiveness of dividend changes to permanent earnings changes and, empirically, the constancy of coefficients.<sup>10</sup>

In economic terms, the growth rate for dividends,  $g(t)$ , equals  $\alpha r(t)$  which is the usual expression for the deterministic growth rate of dividends where  $r(t) \equiv 1 - D(t)/E^m(t-1)$  is the retention rate (in terms of permanent earnings) at time  $t$  and  $\alpha$  is defined as the aggregate cost of capital. That is, specification of the deterministic component  $g(t)$  reflects the standard textbook proposition that if the current payout is high relative to permanent earnings and therefore the retention rate  $r(t)$  is low, then

dividends per share will be expected to grow more slowly than if the current payout were lower and the retention rate were correspondingly higher. The rest of the terms on the right-hand side of (5) describe the deviation of the growth in dividends from this normal rate.

The second term in (5), which is multiplied by  $\lambda$ , captures Lintner's stylized fact (iii) that managers will change dividends away from the anticipated path in response to an unanticipated change in permanent earnings,  $\{\log[E^m(t)/E^m(t-1)] - m(t-1)\}$ . The third term, which is multiplied by  $\gamma$ , is the "error correction" component which drags short-run dividends toward their long-run steady-state payout ratio, thus capturing Lintner's stylized fact (i). The value of  $\gamma$ , which should be positive, measures the average speed of convergence of the payout ratio to its target.

The a priori reasons for choosing the lag specification in (3) and (5), in which an unanticipated change in permanent earnings from time  $t-1$  to time  $t$  causes a dividend change in the interval  $(t, t+1)$ , are as follows: first, an unanticipated change in permanent earnings, by definition, cannot be known until it happens, so any reaction in dividends to such a change must occur at the same time or later. Unlike delays in the reaction of speculative prices to new information, there are no arbitrage opportunities created by managers if they delay changing dividends in response to new information. Second, although firms usually declare dividends once a quarter, many firms only make significant changes at the end of their fiscal year. Third, even if individual firms' managers did react instantaneously, the reaction in aggregate dividends will appear to be lagged because of different announcement dates and different speeds of reaction across firms.

In responding to an unanticipated change in permanent earnings, managers



will change dividends in the same direction which implies that  $\lambda$  in (5) should be positive. From stylized fact (iv), managers prefer to avoid reversals in dividends, and it can be established that a partial adjustment policy with  $\lambda < 1$  is optimal if reversals or changes are costly.

### 3. The Dividend Model Expressed as a Regression Equation

In the empirical studies of both Lintner's model and subsequent dividend models based on his original formulation, the equations corresponding to our (5) are treated as regression equations. We too assume that equation (5) is both a structural equation and a causal equation because our view of the economic process is that an unanticipated change in permanent earnings causes a predictable change in next period's dividends, and not the reverse.

Of course, in a complete general equilibrium model, dividend changes and intrinsic value changes, along with other quantities and prices, are jointly endogenous. However, insofar as the bivariate series of dividend changes and intrinsic value changes is concerned, there are persuasive grounds for treating the latter as a proper predetermined endogenous variable, particularly when the discount rate  $\alpha$  is assumed constant.

As already noted, there are no important legal or accounting constraints on dividend policy, and hence, managers have almost complete discretion and control over the choice of dividend policy. If, however, managers are not irrational, then they will, at least, choose a dividend policy which is feasible in both the short and long runs. Such "feasible" policies must satisfy the constraint that the discounted value of expected future dividends per share is equal to the discounted value of expected future net cash flows as given by (1). Because managers set dividend policy, this constraint is properly specified in terms of their probability assessments. Hence, from (2), it follows that a rational dividend policy must satisfy:<sup>11</sup>

$$E_t^m \left\{ \int_t^{\infty} D(s) e^{-\alpha(s-t)} ds \right\} = E^m(t)/\alpha \quad (6) \quad .$$

This constraint on dividend choice is very much like the intertemporal budget constraint on consumption choice in the basic lifetime consumption decision problem for an individual. In this analogy, the capitalized permanent earnings of the market portfolio,  $E^m(t)/\alpha$ , correspond to capitalized permanent income or wealth of the individual, and the dividend policy of firms corresponds to the consumption policy of the individual. Just as there are an uncountable number of consumption plans which satisfy the consumer's budget constraint for a given amount of wealth, so there are an uncountable number of distinct dividend policies which satisfy (6) for a given value of permanent earnings. Hence, like consumers in selecting their plans, managers have a great deal of latitude in their choice of dividend policy.<sup>12</sup> The fact that individual firms pursue dividend policies which are vastly different from one another is empirical evidence consistent with this view.

Pursuing the analogy further: if because of an exogeneous event (for example, a change in preferences), a consumer changes his or her planned pattern of consumption, then it surely does not follow from the budget constraint that this change in the future time path of his consumption will cause his current wealth to change. Just so, it does not follow from (6) that a change in dividend policy by managers will cause a change in their current assessments of permanent earnings.<sup>13</sup> For a fixed discount rate,  $\alpha$ , it does however follow from (6) that an unanticipated change in permanent earnings must necessarily cause a change in expected future dividends, and this is so, for the same feasibility reason, that with a constant discount rate, an unanticipated change in a consumer's wealth must necessarily cause a change in his or her planned future consumption.

The direction of causation between unanticipated changes in permanent earnings and changes in subsequent dividends posited in our model is, thus, consistent with the direction of causation between changes in current wealth and changes in future consumption that is normally assumed for the life cycle model in a fixed-discount rate world.<sup>14</sup>

In making the case for causality in equation (5), we are not unaware of the possibility that there are other exogeneous variables which may cause managers to change dividends. If this is the case and if, further, these variables are correlated with unanticipated changes in current permanent earnings, then, of course, equation (5) is flawed as a causal equation. If, however, managers are rational predictors of permanent earnings, then an unanticipated change in permanent earnings this period will be uncorrelated with all variables which are observable prior to this period (including both past dividends and permanent earnings). It therefore must also be uncorrelated with all future unanticipated changes in permanent earnings. Thus, if there are other exogeneous variables which explain next period's change in dividends, it seems unlikely that they would be correlated with this period's unanticipated change in permanent earnings. Hence, the assumption that equation (5) is a proper regression equation is likely to be robust with respect to other "missing" explanatory variables.

This property of rationally-predicted permanent earnings together with the lagged structure of equation (5) may perhaps at first suggest that the causality issue can be resolved empirically by applying an appropriate version of the Granger-Sims test of causality. A careful review of this possibility will, however, lead to the well-known identification problem that statistical tests alone are not sufficient to establish causality, and that ultimately

this issue can only be resolved by a priori economic reasoning (cf. Zellner (1979)).

#### 4. A Reduced Form for the Dynamic Model

Although we have proposed equation (5) for aggregate dividend dynamics as both a structural and a regression equation, it cannot be estimated in its current form because management assessments of changes in firms' permanent earnings are not observable. In this section, we add the necessary further specification to estimate the model.

If managers are rational forecasters and the market is reasonably efficient, then the market's estimate of a firm's intrinsic value should on average be equal to the intrinsic value estimate made by that firm's management. We therefore assume that the discounted value  $V^M(t)$  of the expected future aggregate net cash flows of all firms per market share, as estimated from the market's information set, is equal to the aggregate sum of the intrinsic values where the intrinsic value of each firm is estimated from the information set of that firm's management.<sup>15</sup> This market efficiency condition can be written as:

$$V^M(t) = V^m(t) \quad \text{for all } t \quad (7)$$

where  $V^M(t)$  is given by (1) with  $\epsilon_t = \epsilon_t^M$ .

We further assume that the stock market price is equal to its intrinsic value, i.e., there are no speculative bubbles. From this assumption and (7), we can write the cum dividend price of a share of the market portfolio at time  $t$ :<sup>16</sup>

$$P_c(t) = \epsilon_t^M \left\{ \int_t^\infty e^{-\alpha(s-t)} [\Pi(s) - I(s)] ds \right\} / N(t) \quad (8)$$

Using the market efficiency condition (7) and the definition of permanent earnings in (2), we can rewrite (8) as:

$$P_c(t) = E^m(t) / \alpha \quad (9)$$

If, as we have assumed, the expected real rate of return on the market,  $\alpha$ , is a positive constant, it follows from (9) that the percentage change in stock market price is equal to the percentage change in managers' assessment of permanent earnings. Substituting for  $E^m(t)$  from (9) and splitting the cum-dividend stock price change  $P_c(t)/P_c(t-1)$  into its two component parts--the ex dividend change  $P(t)/P(t-1)$  and the dividend yield  $D(t)/P(t-1)$ , we rewrite (5) as:

$$\begin{aligned} \log\left[\frac{D(t+1)}{D(t)}\right] &= \left[\alpha - \frac{D(t)}{P(t-1)}\right] + \lambda\left\{\log\left[\frac{P(t) + D(t)}{P(t-1)}\right] - \alpha\right\} \\ &+ \gamma\left[\rho - \log\frac{D(t)}{P(t-1)}\right] + u(t+1) \end{aligned} \quad (10)$$

where  $\rho \equiv \beta + \log \alpha$ . By rearranging terms, we can rewrite (10) as:

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = a_0 + a_1 \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] + a_2 \log\frac{D(t)}{P(t-1)} + u(t+1) \quad (11)$$

where  $a_1 = \lambda$ ;  $a_2 = -\gamma$ ; and  $a_0 = (1 - \lambda)\alpha + \gamma\rho$ .

Note that (9) is not an identity. It is a specification which is valid under the hypothesis that market prices are rational predictors of firms' future net cash flows. If market price is a very noisy estimator of managements' assessments of permanent earnings, then (10) should be a poor predictor of the evolution of dividends. This would be the case if, on the one hand, managers are rational predictors of future cash flows but, on the other, the market price is subject to wide swings caused by waves of irrational investor optimism and pessimism. Alternatively, (10) could be a poor predictor because management dividend decisions in the aggregate are not



well described by behavioral equation (5). Such is the nature of a reduced form equation.

If, however, market price provides a good estimate of managements' assessments of permanent earnings, then (10) should be a good predictor of the dividend process. Such would be the case if managers are rational predictors of future cash flows and the market is efficient. It is, of course, possible for the market to be inefficient and for (10) to be a good predictor of the dividend process. This would be the case if the market is inefficient because it is moved by waves of optimism and pessimism, and managers either rely on market price as their estimate of permanent earnings or they are influenced by the same irrational waves as investors in making their assessments of intrinsic value.<sup>17</sup>

For (11) to be a proper reduced form equation, its right-hand side variables must be predetermined relative to its left-hand side variable. As discussed at length in Section 3, an unanticipated change in permanent earnings this period (over which managers have no control) is an exogeneous variable relative to the change in next period's dividends which managers control almost completely. From (7) and structural equation (9), it therefore follows that an unanticipated change in this period's price is exogeneous relative to next period's dividend change, and hence, (11) is a proper reduced form equation. In this limited sense of a reduced form, an unanticipated change in this period's price "causes" a (predictable) change in next period's dividends.

In specifying (11), our intent is to construct a simple model of the dividend process which nevertheless captures the basic stylized facts of management behavior. We have therefore assumed a simple one-period lagged



adjustment. It is possible that the dividend process may involve higher order lags with different speeds of adjustment, and as already noted, there may be other "missing" variables which enter into the process. As we will show, the empirical conclusions derived from this simple model are likely to be robust with respect to refinements which include such additional variables.

## 5. Model Estimation

To estimate the reduced form equation (11), we use annual data<sup>18</sup> constructed from monthly dividend and price series for the value-weighted NYSE index contained in the Center for Research in Security Prices data set over the period 1926 to 1981.<sup>19</sup> Over the period 1927 to 1980, the discrete time version of (11) estimated by ordinary least squares (OLS) is given by

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{-0.101}{(0.157)} + \frac{0.437}{(0.064)} \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] + \frac{-0.042}{(0.050)} \log\left[\frac{D(t)}{P(t-1)}\right] + u(t+1) \quad (12)$$

$$\bar{R}^2 = 0.47 \quad DW = 1.53$$

In (12),  $D(t)$  refers to aggregate NYSE dividends totalled over year  $t$ , and  $P(t)$  refers to price at the end of year  $t$ . The numbers in parentheses under the coefficients are standard errors, and not  $t$ -statistics. For example, the coefficient of the lagged logarithmic change in price has a standard error of 0.064, and a  $t$ -statistic of 6.83. The coefficient point estimates in (12) indicate that the deviations of real dividend changes from their normal growth rate covary positively and strongly with the previous year's unexpected cum-dividend price changes and negatively with the previous year's dividend yield. The Durbin Watson statistic suggests that there is positive

autocorrelation in the residuals of the OLS fit of (12). Disturbance correlation can arise in various ways. As already noted, our simple model assumes one-period adjustment in dividends by management, whereas longer lags are entirely possible. Further, as we noted earlier, it is possible that the target dividend yield is not literally constant. We suggested that the yield might change if tax rates, the technology of communications and trading, or the mix of institutional and individual ownership, change. It is highly likely that any yield changes induced by these factors, which will show up in the residuals in (12), will be serially correlated. Indeed, any omitted variables which are serially correlated could be a potential source of residual autocorrelation.<sup>20</sup>

In light of the autocorrelation in the residuals of (12) we reestimated (11) using generalized least squares, and the results are:

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{-0.234}{(0.198)} + \frac{0.444}{(0.061)} \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] \quad (13)$$

$$-0.085 \log\left[\frac{D(t)}{P(t-1)}\right] + u'(t+1)$$

$$(0.082)$$

$$\bar{R}^2 = 0.53 \quad DW = 1.83$$

Although the GLS estimate which takes into account the positive autocorrelation appears to have slightly more explanatory power, the results from either the OLS or GLS fits are essentially the same in that they explain about 50 percent of aggregate NYSE real dividend changes. As will be shown in Section 6, the explanatory power of our single-equation aggregate time series model<sup>21</sup> is considerably higher than that of univariate trend-autoregressive models such as that fitted by Shiller (1981a).

The point estimate of 0.44 for the coefficient on the lagged percentage price change is positive, substantial in magnitude, and highly significant. This finding is consistent with the hypothesis that the market price is a good indicator of permanent earnings and that managers systematically change dividends in response to an unanticipated change in permanent earnings.

Further, the estimate of the coefficient on percentage price changes shifts only negligibly from the OLS estimate of 0.437 in (12) to the 0.444 in (13) when GLS is used. If we interpret the GLS transformation as a quasi-differencing operator, and use a logic similar to that of Plosser, Schwert, and White (1982), then the invariance of the coefficient estimate to GLS suggests, in terms of Hausman's (1978) specification test, that there is no simultaneity bias in regressions (12) or (13). That is, the evidence is consistent with lagged price changes "causing" dividend changes, but not vice versa.

Because the coefficient on percentage price changes is both significantly greater than zero and significantly less than one, this finding is also consistent with the Lintner stylized fact that managers smooth dividends by responding in a partial adjustment fashion to an unanticipated change in permanent earnings. The well-established empirical fact that the variation in the percentage change in dividends is significantly less than the variation in the percentage change in prices, might suggest to some that prices are "too volatile." However, the empirical verification in (12) and (13) of the partial adjustment mechanism posited in our model provides an explanation of this well-established fact that is entirely consistent with market price being a rational predictor of future dividends.

The estimated coefficient of the dividend-to-price ratio is negative in both (12) and (13), which is consistent with the hypothesis that this ratio converges to a long-run stationary distribution. The point estimates for the speed of adjustment are rather small which suggests that a substantial period of time is required for the dividend-to-price ratio to converge to its steady-state distribution. Thus, using the  $-0.085$  estimate from (13), a conventional "half-life" calculation shows that it takes more than eight years for the expected value of this ratio to move halfway from its initial value to its expected steady-state value.<sup>22</sup>

To further investigate the extent of synchronization between dividend changes and price changes, we computed the leads and lags of the percentage changes in dividends regressed on percentage changes in prices estimated by the Hannan-efficient procedure, and these are plotted in Figure 1. By inspection, the cross correlation at the lag in price change of one year specified in our model dominates that at all other leads and lags. In even a reasonably efficient market, one would not expect lagged variables of any sort to have meaningful predictive power for future price changes. It is therefore not surprising that changes in dividends are not significantly correlated with subsequent changes in price. The modest positive correlation between contemporaneous dividend and price changes is, of course, consistent with an efficient market and is perhaps suggestive of a mild information or announcement effect for dividend and price changes at the aggregate level. Indeed, we do find about an 8 percent correlation between contemporaneous (i.e., year  $t + 1$ ) unanticipated price changes and the residuals from our regression (13). As noted in the discussion of the informational content of dividends in the Introduction, it is difficult to identify an announcement

date for aggregate dividends in a meaningful way. Moreover, what is perceived to be contemporaneous correlation between dividend and price changes over the coarse grid of annual data may simply turn out to be lagged price changes explaining subsequent dividend changes when examined under the finer grids of quarterly or monthly data. Thus, an 8 percent correlation is likely to be a significant overstatement of the announcement effect of aggregate dividends.

Unlike for speculative price changes, the Efficient Market Hypothesis does not rule out the change in dividends this period being predicted by variables which are observable prior to this period. Nevertheless, if the posited economic process underlying the specification of (11) is a reasonably accurate description of reality, then lagged price changes much beyond the one-year lag specified in (11) might be expected to exhibit relatively little explanatory power in forecasting this period's dividend change. If an unanticipated change in permanent earnings causes managers to change dividends, then strict rational forecasting would seem to dictate that their decision be based on their most recent assessment of permanent earnings, and hence earlier revisions in those assessments should have little impact on the change in dividends. In attempting to smooth the time path of dividends, it is possible that managers would choose to change dividends in response to changes in a moving average or distributed lag of unanticipated permanent earnings changes over an extended past history. Such behavior would create a dependency between the current change in dividends and lagged price changes of all orders. Because these averaging techniques embody much "stale" information, it would appear that this approach to dividend smoothing leads to an inefficient use of the available information. If instead, managers change the dividend in a partial adjustment response to the most recent unanticipated



change in permanent earnings, they can use the most up-to-date information. The partial dividend adjustment will be appropriate in a wide variety of contexts in which the policy problem facing managers does not admit a certainty-equivalence solution.

Even if managers forecast this rationally, there will still be some lag between a change in permanent earnings and the change it induces in subsequent dividends. As indicated in the discussion surrounding the specification of equation (5), at the level of aggregate dividends, it is unlikely that this "minimal" lag is much shorter than a year.<sup>23</sup> The correlations between the change in dividends and lagged price changes displayed in Figure 1 are therefore consistent with this view of rational forecasting by managers. As a further, more quantitative test, (11) was reestimated with five years of lagged unexpected price changes as additional variables. None of these additional lagged variables had a coefficient point estimate more than one standard error from zero, and the  $F$  statistic for their inclusion is 0.361, which has a  $p$  value of 0.872.

If these empirical results had turned out differently, they would neither imply an arbitrage opportunity in the stock market nor an inefficiency in the allocation of capital. We need hardly mention again that managers have great latitude in their selection of dividend policies including the option to choose ones which are not based on the most up-to-date information. It is, however, reassuring for the overall validity of our model that the data tend to support such "superrational" forecasting behavior by managers even in the relatively unimportant area of dividend policy.

With the exception of contemporaneous changes in other speculative prices, it is a well-established empirical fact that there are few, if any, observable variables which exhibit high contemporaneous correlation with changes in

aggregate stock prices. It is, therefore, rather unlikely that the change in stock price is merely serving in (11) as a proxy for some other observable variables which, if included, would cause the significance of the coefficient on the price change to disappear.

We have investigated whether the dividend response to stock price changes is symmetric with respect to negative and positive price changes. The point estimate of this elasticity with respect to the twenty negative annual price change observations in our sample is 0.597, while it is 0.271 with respect to the positive price changes. While an asymmetry of dividend response to negative and positive price changes could be readily explained in a more complete dividend model, the small number of negative price change observations results in a standard error for the difference in point estimates that is so large that equality of the elasticity coefficients could not be rejected. In addition, two of the negative price change observations are those for the 1929 and 1974 "crashes" in which the constant discount rate assumption, and possibly constancy of the model itself, is strained.

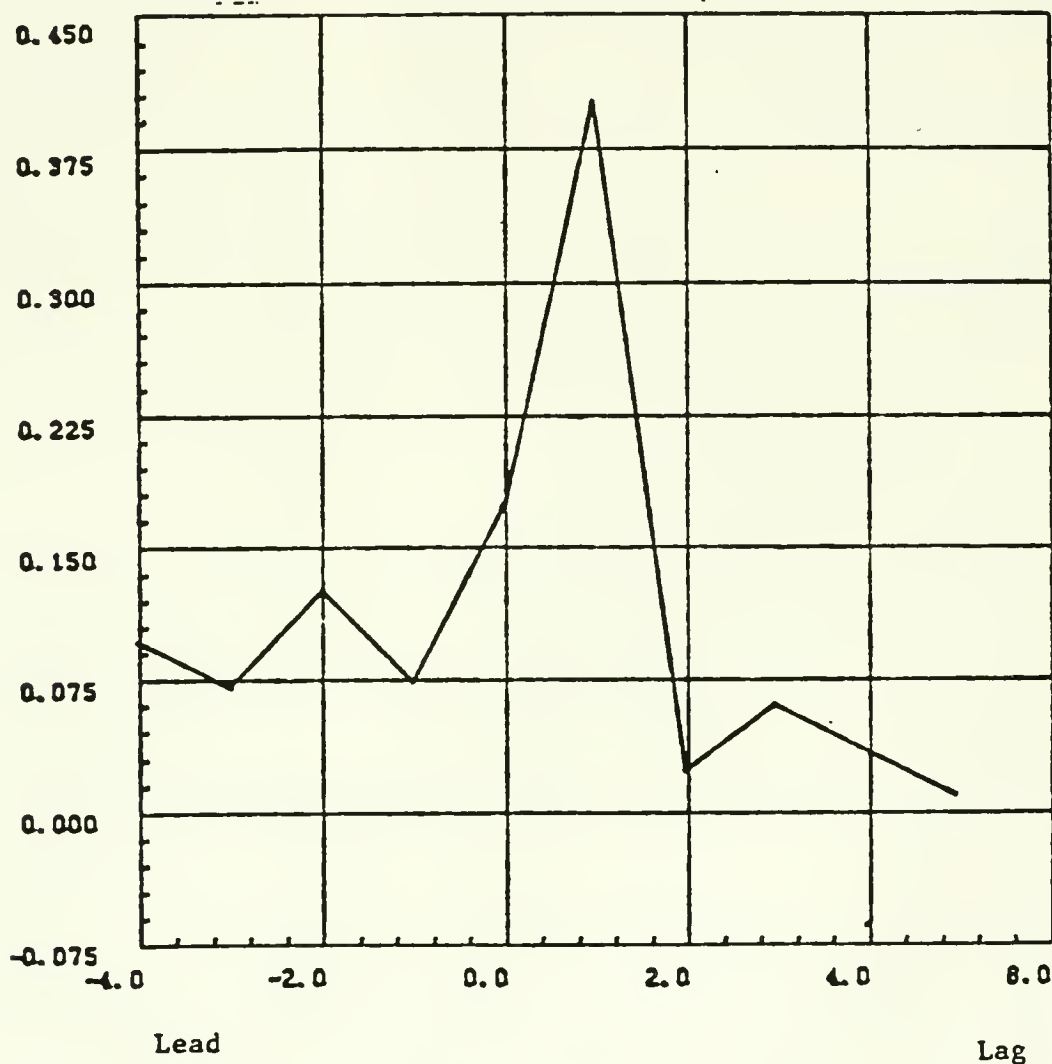
We have also used recursive regression techniques to check for parameter stability in (5) over our sample period. There does appear to be some secular decline in the elasticity variable  $\lambda$  and the regressivity coefficient  $\gamma$ --most markedly so after World War II--though only the latter is significant. As already suggested, the factors which could historically have given rise to the importance of dividend policy may themselves have diminished in significance, so a secular trend would not be surprising. On this score, a new set of survey data regarding the importance of corporate dividend policies today would be particularly pertinent.

In summary, our main empirical results are that (i) past (or possibly



Figure 1: Lead and lag structure of deviations in annual percentage real dividend changes, around their expected growth rate,<sup>1</sup> on percentage unexpected real price changes,<sup>1</sup> for the NYSE value-weighted companies over the period 1927-1980, computed using the Hannan-efficient procedure.

Point Estimate  
of  $\hat{a}_1$  in (\*)



$$\text{Equation: } \log\left[\frac{D(t+j)}{D(t+j-1)}\right] + \frac{D(t+j-1)}{P(t+j-2)} = a_0 + a_1 \log\left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-2)}\right] \quad (*)$$

$$j = -4, \dots, 0, \dots, +6$$

<sup>1</sup>Dividends,  $D(t)$ , are defined as year  $t$  cash dividend payments for all NYSE companies, and the index "price"  $P(t)$  is the end of year  $t$  value weighted index. The dividend-price ratio,  $D(t+j-1)/P(t+j-1)$  is, up to a constant, the expected rate of growth of dividends and prices in year  $t+j$ .

contemporaneous) changes in stock market prices explain a significant portion of the change in aggregate dividends, and (ii) the (partial) elasticity of the dividend response to a change in price is positive and significantly less than one. Further, there are reasonable grounds for believing that these results will be robust to more refined versions of this dividend model. Armed with these results, we show, in Appendix A, that in this model where rationally-determined prices, and thus rationally-determined dividends, do not have a nondegenerate steady-state distribution (with or without a trend), the dividend-to-price-ratio follows a stationary process. We thus directly establish that even if price levels and dividend levels follow integrated stochastic processes, as they seem to do empirically (cf., Kleidon (1983)), the target payout ratio can satisfy the steady-state properties assumed of it in the reduced form of our error correction model (11).

## 6. Further Discussion, and Comparison of Our Model With Others in the Literature

In this section, we compare the fit of the model of aggregate dividends which has been developed in previous sections with that of the "final form" trend-autoregressive model employed by Shiller (1981a) to describe aggregate dividends. We show that our model fits considerably better than does the trend-autoregressive model, and that lagged dividends explain little, if any, of the variation in aggregate dividends once prior-period-stock-price-changes are taken into account. We also argue that the characteristics of the observed distributions of dividend and stock price changes reported by Shiller can be readily interpreted within the context of our model.

In Section 6.2, the explanatory power of our model of aggregate dividend behavior, in which stock price changes are used to measure changes in firms' permanent earnings, is also compared with models such as those fitted by Lintner (1956), Brittain (1966), and Fama and Babiak (1968) in which dividend movements are explained by accounting earnings changes. We find that our model fits at least as well as accounting-earnings-based models, even when contemporaneous accounting earnings data is used in the latter while only prior period stock price changes are used in the former.

Finally, we briefly discuss, in Section 6.3, similarities in the fit of our model to gross and net dividends.

## 6.1 A Trend-Autoregressive Model for Aggregate Dividends?

In his most recent published evidence, Shiller (1983) reports that "If  $\log D(t)$  is regressed on  $\log D(t-1)$ , a constant and a linear time trend for 1872 to 1978, the coefficient of  $\log D(t-1)$  is 0.807, with an estimated standard error of 0.058," implying that "log dividends would always be expected to return half way to the trend line in three years," (p. 237). We repeated essentially the same OLS regression on our data set, and the results are:

$$\Delta \log D(t) = \begin{matrix} 2.492 \\ (0.890) \end{matrix} - \begin{matrix} 0.249 \log D(t) \\ (0.089) \end{matrix} + \begin{matrix} 0.004 t \\ (0.002) \end{matrix} + u(t) \quad (14)$$

$$\bar{R}^2 = 0.130 \quad DW = 1.495$$

Because the left side of (14) is the change in  $\log D(t)$ , the comparable autoregressive coefficient is  $(1 - 0.249) = 0.751$  which is rather close to Shiller's 0.807 estimate. By the standards of a conventional t-test, the coefficients in both samples are significantly less than 1 with a -2.80 t-statistic in our sample and a -3.33 t-statistic in his. The t-statistic for the trend coefficient in (14) is 2.00. This finding serves to confirm our belief that the important empirical results derived from our 1926-1981 data set are not likely to be significantly altered if fit to the longer 1871-1979 data set used by Shiller in his analysis of aggregate dividend and stock price behavior.

Because the Durbin Watson statistic suggests positively autocorrelated residuals, the lagged endogeneous variable in (14) may cause the OLS coefficients to be biased. We therefore reestimated (14) using a GLS iterative technique,<sup>24</sup> and the results are:

$$\Delta \log D(t) = \frac{5.225}{(1.211)} - \frac{0.524}{(0.121)} \log D(t) + \frac{0.009}{(0.003)} t + u(t) \quad (15)$$

$$\bar{R}^2 = 0.243$$

$$DW = 1.85$$

Hence, when the OLS specification of Shiller's autoregressive model for dividends is correctly adjusted for autocorrelation, its measured explanatory power almost doubles. That is, only about half of the total 24 percent explanation of the variation in dividend changes can be attributed to the lagged dividend and time trend variables. The other half is attributable to the time series model of the disturbances or "unknown variables" in the regression.<sup>25</sup>

Given the apparent statistical significance of the coefficients in (14) and (15), it is perhaps tempting to some to conclude that dividends follow an autoregressive process which approaches a steady-state distribution (possibly around a positive trend). Such a conclusion, if true, has far-reaching implications for the whole of financial economics. For example, if (14) fully described the "true" dividend process (i.e.,  $\bar{R}^2 = 1$ ), then this result, along with even a casual inspection of the stock return time series, would surely imply that stock prices are "too volatile." If, as implied by (15), dividends were to regress over 90 percent of the way to their deterministic trend line within one Presidential term, then uncertainty about the future path of dividends would be rather unimportant, and therefore, rational stock prices should exhibit trivial fluctuations. Because equities are the residual claims of the private sector, variations in their returns are "blown up" reflections of the uncertainties about the whole economy. If stock returns should have small variations, then the fluctuations in the economy should be even

smaller. It would therefore seem that in such an environment we economists could safely neglect such uncertainties in the specification of our macroeconomic models. While perhaps an appealing hypothesis, the real world is not this way as further analysis of (14) and (15) will clearly indicate.

The fit of the autoregressive model (14) and (15) is rather poor, with half of the explanatory power of (15) represented by unspecified variables. With respect to a different regression on similar data, Shiller (1981a, p. 433) gives one possible explanation for low  $R^2$ : Namely, "...regression tests are not insensitive to data misalignment. Such low  $R^2$  might be the result of dividend or commodity price index data errors." Although we agree that such data errors can be a source for lower  $R^2$ , our explanation is simply that the autoregressive process posited in (14) and (15) is not an accurate specification of the dividend process.

Motivated by the analysis of our model of the dividend process, we add the one-year lagged unanticipated change in the log of stock price to the specification of (15). By the same iterative GLS procedure used in (15), the results are:

$$\begin{aligned} \Delta \log D(t) = & \frac{2.107}{(0.918)} + \frac{0.347}{(0.064)} \log \left[ \frac{P(t)}{P(t-1)} + \frac{D(t)}{P(t-1)} \right] \\ & - \frac{0.213}{(0.092)} \log D(t) + \frac{0.004}{(0.002)} t + u(t) \end{aligned} \quad (16)$$

$$\bar{R}^2 = 0.473$$

$$DW = 1.755$$

By inspection, the addition of the previous year's unexpected price change in (16) doubles the explanatory power of (15). This measured increase in  $\bar{R}^2$  greatly understates the impact of this added variable because, in addition to increasing  $\bar{R}^2$  by 100 percent, it also virtually eliminates the explanatory



power of the remaining unspecified variables whose effects are captured by the GLS procedure.<sup>26</sup> We would also note that by adding the log price change variable, the absolute magnitudes of both the  $\log D(t)$  and time trend coefficients are cut in half.

To further explore the relative importance of the specified variables, equation (16) was reestimated: first, with the time trend deleted, and second, with both the time trend and  $\log D(t)$  removed. The results are:

$$\begin{aligned} \Delta \log D(t) = & \frac{0.566}{(0.564)} + \frac{0.388}{(0.064)} \log \left[ \frac{P(t)}{P(t-1)} + \frac{D(t)}{P(t-1)} \right] \\ & - \frac{0.055}{(0.053)} \log D(t) + u(t) \end{aligned} \quad (17a)$$

$$\bar{R}^2 = 0.437 \quad DW = 1.814$$

and

$$\Delta \log D(t) = \frac{-0.014}{(0.016)} + \frac{0.402}{(0.063)} \log \left[ \frac{P(t)}{P(t-1)} + \frac{D(t)}{P(t-1)} \right] + u(t) \quad (17b)$$

$$\bar{R}^2 = 0.435 \quad DW = 1.82$$

Comparing (16) with (17a), the elimination of the time trend variable causes only a modest reduction in  $\bar{R}^2$ , and it has little effect on the estimated coefficient of the log price change. However, by eliminating the time trend variable, the magnitude of the regressive coefficient on  $\log D(t)$  falls by 75 percent, and with a  $p$  value equal to 0.309, it is not statistically significant. It would appear that there is a strong interaction between  $\log D(t)$  and the time trend which together with the GLS iterative procedure is responsible for the significant coefficients in (15). If either variable is removed, then the magnitude and the statistical significance of the coefficient of the remaining variable are both nil. In this light, it is not

surprising that the elimination of  $\log D(t)$  as a variable in (17b) has no effect on either the  $\bar{R}^2$  of that equation or the coefficient of log price change.

Unless the log price change can be "explained" by some distributed lag of past dividends (which, as an empirical matter, it cannot), then it surely belongs in the specification of the dividend process. Because it alone accounts for over 90 percent of the explanatory power of (16), its omission from (14) and (15) is rather important. In sharp contrast, the elimination of either the time trend or the  $\log D(t)$  variables has no significant effect on the fit. Hence, unless there are strong a priori economic reasons to believe that these variables belong in the specification of the dividend process, there appears to be no valid empirical reason for their inclusion.

In our model of the dividend process, there is no role for a time trend. Its inclusion produces an insignificant coefficient, and actually causes the corrected OLS and GLS  $R^2$ 's in (12) and (13) to fall. Our model would, however, predict that changes in log dividends are related to  $\log D(t)$  through the dividend-to-price ratio,  $\log[D(t)/P(t-1)]$ . Although not explicit in our simple model, it is possible that lagged changes in log dividends could explain part of the adjustment process used by managers to decide upon subsequent dividend changes. If this were the case, then  $\log D(t)$  may be a proxy for these lagged changes. The inclusion of such lagged dividend changes would in no substantive way change the conclusions derived for the dividend and rational stock price processes in Section 5. To investigate this possibility, we reestimate equation (13) with the addition of  $\log D(t)$  and the fitted results are given by:



$$\begin{aligned} \log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} &= \frac{1.550}{(1.923)} + \frac{0.441}{(0.061)} \log\left[\frac{P(t)}{P(t-1)} + \frac{D(t)}{P(t-1)}\right] \\ &- \frac{0.247}{(0.267)} \log\left[\frac{D(t)}{P(t-1)}\right] - \frac{0.220}{(0.256)} \log D(t) + u(t+1) \end{aligned} \quad (18)$$

$$\bar{R}^2 = 0.532 \quad DW = 1.88$$

The F statistic for  $\log D(t)$  in (18) is 3.23 which is insignificant at the 5 percent level. If  $\log D(t-1)$  is added to (18), its estimated coefficient is -0.135 while that of  $\log D(t)$  becomes 0.064. While this result suggests that  $\log D(t)$  in (22) may be a proxy for  $\log[D(t)/D(t-1)]$ , the coefficient of  $\log D(t-1)$  is also statistically insignificant. As expected, the addition of these lagged dividend variables in (18) had no effect on the point estimate of the coefficient of lagged log price change.<sup>27</sup>

In summary, adding the trend and lagged dividend variables of the Shiller model to our specification does little to improve the explanatory power of our model in terms of  $\bar{R}^2$ . Moreover, the estimated coefficients of these "added" variables are statistically insignificant. The other side of this result is that the addition of the variables from our model to the autoregressive specification (14) substantially increases the explanatory power of that model. As these variables are added, however, the statistical significance of the autoregressive variables is reduced. This result is perhaps surprising because it is a common belief that the dividend time series is quite smooth by comparison with the price series. Thus, a distributed lag of past dividends together with a time trend might be expected to do a better job than stock price changes in explaining subsequent dividend changes, almost independently

of the "true" economic specification.<sup>28</sup>

Although there appears to be no significant empirical evidence for regressivity in the time series of dividends, the lack of such evidence does not disprove the hypothesis that dividends have a stationary distribution around a deterministic trend. As discussed at length in Section 3, the resolution of such an issue must ultimately come from economic reasoning. As Shiller (1983, p. 236) has noted on the specification issue, "Of course, we do not literally believe with certainty all the assumptions in the model which are the basis of testing. I did not intend to assert in the paper that I knew dividends were indeed stationary around the historical trend." We surely echo this view with respect to the theoretical assumptions underlying our own empirical model. Nevertheless, unlike our model's assumptions, the assumption that dividends follow a stationary process with a trend does not appear to have any nonstatistical foundation. There appears to be little, if any, theoretical structure to support this assumption. In particular, there is neither an oral nor a written tradition in the financial economics literature that assumes dividends and rational stock prices have stationary distributions.<sup>29</sup>

One could, of course, revive Malthus, or the more contemporary "limits-to-growth" view of economics to justify the assumption of a steady-state distribution for the levels of dividends and prices. This theory, however, also rules out an exponential growth trend in these levels. Refitting equation (14) without the trend, we have that the OLS estimate is given by:

$$\Delta \log D(t) = 0.802 - 0.076 \log D(t-1) + u(t) \quad (19)$$

(0.576) (0.055)

$$\bar{R}^2 = 0.017 \quad DW = 1.576 \quad .$$

By inspection of (19), it would appear that at least in the dividend series, there is no evidence at all to support this "zero-growth" model.

Notable perhaps by its absence from this section is any discussion of tests of stationarity of stock price levels. We have not directly tested the time series of price changes for evidence of regressivity because the literature is almost uniform in failing to find any lagged variables which have much power in forecasting future stock price changes. Moreover, using autocorrelation and Dickey-Fuller (1979) tests, Kleidon (1983) finds that neither the arithmetic nor the geometric Brownian motion models can be rejected against the trend model for the S&P 500 annual composite index over the period 1926-1979. We would expect that these same results would obtain for our data set.

Shiller (1981a, p. 432-433) does report that the dividend-to-price ratio appears to forecast next period's holding period returns. We replicated this result on our data set, and found, as did he, that the  $R^2$  is about 0.06. As Shiller himself stresses, such regression tests are sensitive to data problems, and such problems could easily explain the positive relation between stock returns and lagged dividend yield.<sup>30</sup> In addition, Miller and Scholes (1982) show that holding period returns for individual firms can be forecasted just as accurately using stock price inverses as they can using dividend yields, which suggests that the numerator of the dividend-price ratio does not play an important role in its predictive power for future stock price changes. Of course, forecastability of discrete period returns is compatible with market efficiency if it reflects no more than forecastable variations in expectations of those returns. Forecastable changes in expected returns probably account for only a small amount of the variation in holding period returns however, just as Shiller's  $R^2$  statistic suggests.

## 6.2 The Distribution of Rational Stock Prices and Dividends In Our Model

As we show in Appendix A, given our dividend model and empirical results, the dividend-to-price ratio has a stationary distribution with a finite variance. It follows that for  $T$  large,  $\text{Var}(\log[P(T)/P(0)])$  will be proportional to  $T$ . Hence, for large  $T$ , the cumulative dynamics for  $P$  can be well-approximated as having come from a geometric Brownian motion with an instantaneous expected rate of growth equal to  $(\alpha - \rho)$  where  $\rho$  is the expected "long-run" dividend-to-price ratio computed from the steady-state distribution for  $D/P$ . In this same sense, the asymptotic process for dividends will also be a geometric Brownian motion with  $\text{Var}(\log[D(T)/D(0)])$  proportional to  $T$ .

In a reply to a comment on his work by Basil Copeland, Shiller (1983, p. 237) notes that "Even if we assumed log dividends were a random walk with trend with independent increments, stock prices still would show too much volatility." As he correctly points out, if  $D(t+1)/D(t)$  is independent of  $D(t'+1)/D(t')$  for  $t \neq t'$ , then the current dividend will be proportional to the current price (i.e., in our notation,  $D(t) = \rho P(t)$ ). Hence, in such a model, the variance of logarithmic dividend changes will equal the variance of logarithmic price changes. Shiller goes on to report that for his Standard and Poor data set from 1871-1979, the sample standard deviation for log dividend changes is 0.127, whereas the sample standard deviation for log prices changes is 0.176. Because the ratio of sample variances of 1.93 is significant at the 1 percent level, he concludes that prices are too volatile to be consistent with this model. In our much-shorter 1926-1981 sample period, the standard deviation of dividend changes is virtually the same as in his sample (0.124), but the standard deviation of

price changes is higher (0.203) which leads to a larger sample variance ratio of 2.64. We therefore agree with Shiller's conclusion, although our description would be that "the sample variations in dividend changes are too small to be consistent with this model."

" $D(t) = \rho F(t)$ " is the extreme polar case of our model where managers do not attempt to smooth dividends at all, and fully and immediately adjust dividends to reflect unanticipated changes in permanent earnings. If the "short-run" instantaneous variance rate for logarithmic price changes is  $\sigma^2$ , then in our model the corresponding short-run variance rate for dividends is  $\lambda^2 \sigma^2$ . If managers fully adjust dividends in response to unanticipated changes in permanent earnings, then  $\lambda = 1$ , and the variance of dividend changes and price changes will be the same for all time intervals. If, as empirically appears to be the case, managers smooth the dividend time path by making partial adjustment responses, then  $\lambda < 1$ , and the variance of dividend changes will be strictly smaller than the variance of price changes. If, indeed, our model completely explained the process for dividend changes, then the coefficient estimate in (13) of 0.44 for  $\lambda$  would imply that the ratio of the variances of annual log price changes to log dividend changes would exceed 5. Because the model explains only about 50 percent of the variation in dividends, the actual ratio is reduced to 2.64.

As is generally true of "smoothed" processes which are constrained to converge to a more-variable process, the variance rate of the percentage change in dividends increases as the interval over which it is computed is increased. It is, however, also the case that for  $\lambda < 1$ :

$$\frac{\text{Var}(\log[D(T)/D(0)])}{T} < \frac{\text{Var}(\log[P(T)/P(0)])}{T}$$



for any interval  $T$  and equality holds only in the limit as  $T \rightarrow \infty$ . Our model of rationally-determined prices and rationally-determined dividends, therefore, predicts that the variance rate of logarithmic dividend changes will always be smaller (and, at least for annual or shorter intervals, considerably smaller) than the variance rate of logarithmic price changes. It is, hence, reassuring to find this prediction confirmed by Shiller's statistics which are based on a considerably longer sample period than our own.

In his 1981a paper (p. 428), Shiller also discusses the higher-order moment properties of the stock price and dividend processes with a focus on the relation between infrequent arrivals of important information and the often-observed high-kurtosis (or "fat tail") sample characteristic of stock price changes. He demonstrates this relation by an illustrative example where dividends are taken to be independently and identically distributed. To capture the effect on stock price of infrequent arrivals of important information, he assumes that at each time  $t$ , with probability  $1/n$ , the market is told the current dividend level and with probability  $(n - 1)/n$ , the market has no information about current or future dividends. In this example, the kurtosis of the stock price change is shown to be  $n$  times greater than the kurtosis of the normal distribution posited for dividends. Our model, however, predicts the exact opposite result: Namely, dividend changes should exhibit relatively higher kurtosis than stock price changes. That is, although the variance of dividend changes is smaller than for stock price changes, the time series of dividends should contain more relatively small changes and more relatively large changes than the corresponding time series of stock prices.

The "lumpy" arrival of information (which may, in fact, cause the sample distribution of stock price changes to have fatter tails than a normal distribution) is not the source of this prediction about dividend changes. Instead, it comes as a result of managers smoothing the time path of dividends. To illustrate this point, we use an example which is very much like Shiller's information example.

Suppose that unanticipated logarithmic changes in stock price are serially independent and identically distributed. Suppose further, that managers smooth the time path of dividends according to the following rule: At each time  $t$ , with probability  $1/n$ , they change the dividend to fully reflect the unanticipated change in stock price (i.e.,  $\log[D(t)/D(t-1)] = \log[P(t)/P(t-1)]$ ) and with probability  $(n-1)/n$ , they change the dividend to equal its expected long-run normal growth rate (i.e.,  $\log[D(t)/D(t-1)] = g$ ). As expected for smoothed processes generally, in this example, the variance of dividend changes,  $\sigma^2/n$ , is smaller than the variance of stock price changes,  $\sigma^2$ . If  $m_4$  denotes the fourth central moment of the stock price change distribution, then the fourth moment of the dividend change distribution is  $m_4/n$ . Hence, the kurtosis of the dividend change process,  $(m_4/n)/(\sigma^2/n)^2$ , is simply  $n$  times the kurtosis of the stock price process,  $m_4/\sigma^4$ . Thus, unless managers do not attempt to smooth dividends at all (i.e.,  $n = 1$ ), the kurtosis of the controlled dividend process will always exceed the kurtosis of the (uncontrolled) stock price process. Indeed, the more strongly that managers attempt to smooth dividends (i.e., the larger is  $n$ ), the greater is the relative kurtosis of the dividend process.



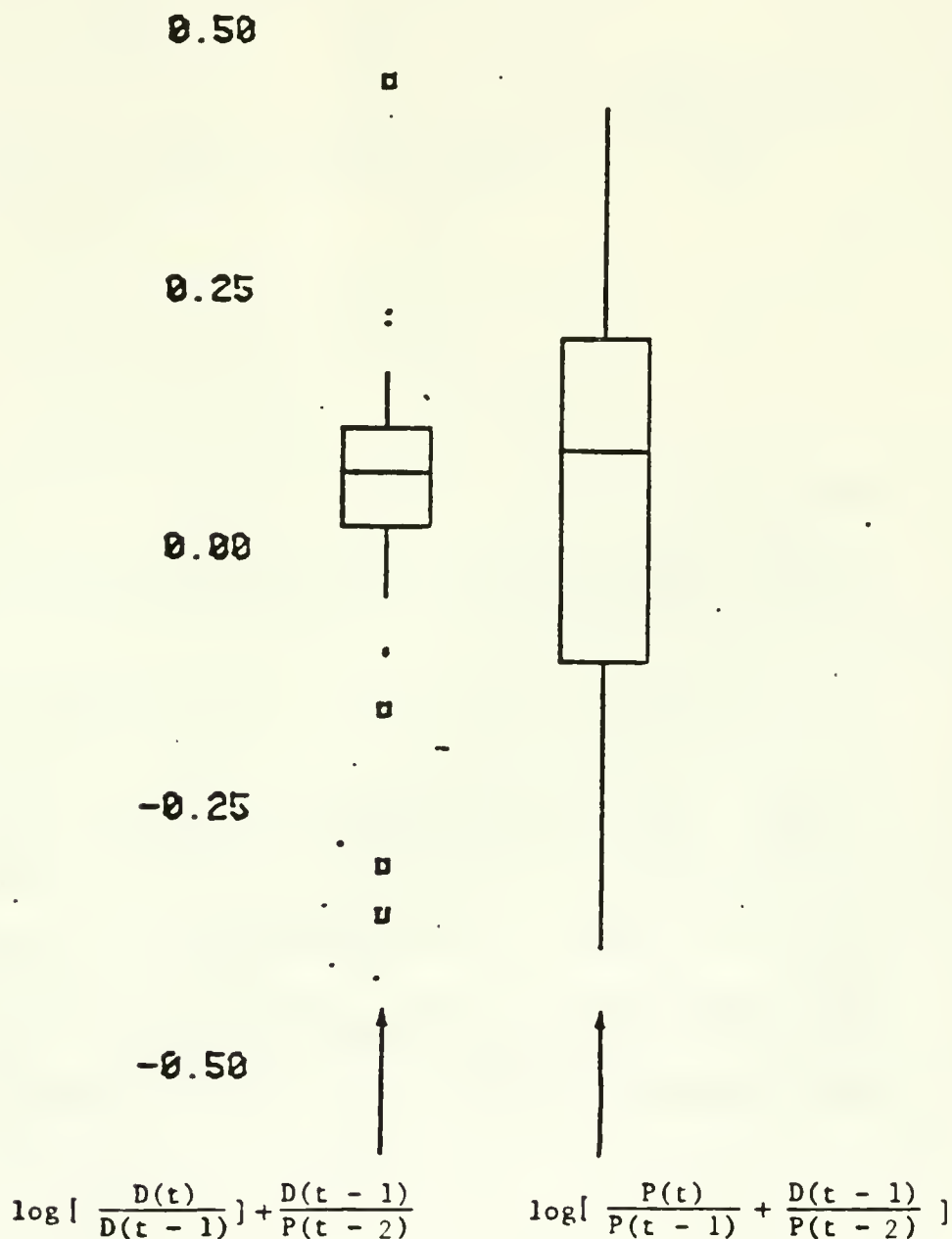
In the light of this result, we estimated the kurtosis of each of the time series as a further empirical check on our model. As predicted, the estimated kurtosis for the annual logarithmic dividend changes, 7.377, is 2.79 times the estimate of 2.648 for the kurtosis of the logarithmic changes in stock prices. As it happens, the sample kurtosis for stock prices is not much different than the kurtosis of 3 for a normal distribution whereas the sample kurtosis for dividend changes is more than two times larger. As further evidence on the relative "fat tails" of the dividend process, the sample distributions of dividend and stock price changes are plotted in Figure 2, using Tukey's (1970) robust statistics. By inspection, these plots of the data are also consistent with the kurtosis prediction of our model.

### 6.3 Stock Price Versus Accounting Earnings as a Measure of Permanent Earnings

In deriving our reduced-form dividend model (11), we adopted the specification (9) that stock prices embody rational predictions of firms' future net cash flows and thus permanent earnings. In past empirical work employing the Lintner model, however, accounting earnings, modified accounting earnings, or cash flow data have been used to measure firms' permanent earnings. Hence, we compare the performance of our model to these alternative measures of permanent earnings.

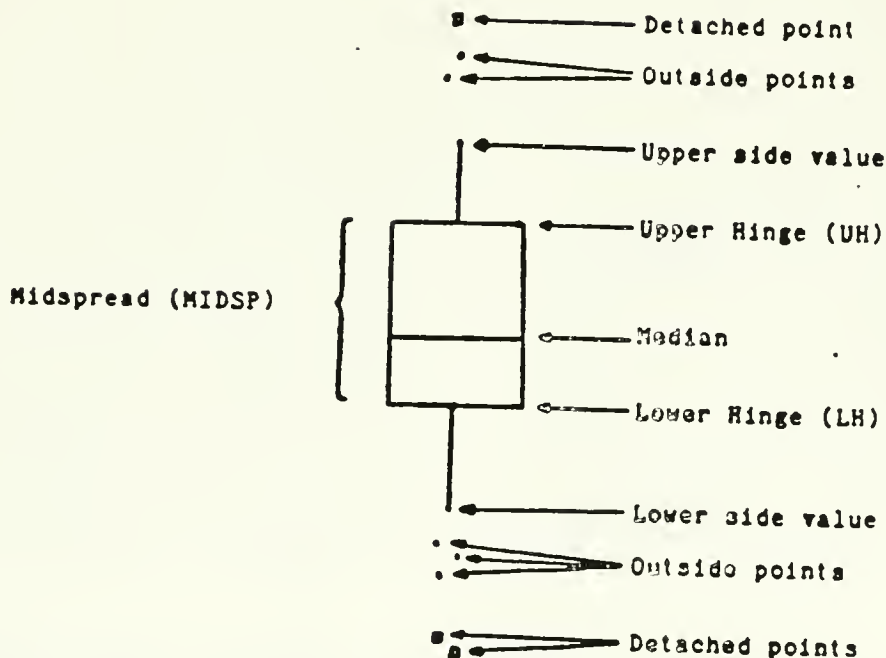
Since the readily available data on aggregate earnings are for the S&P 500 companies, the results reported in this section pertain to this index. The fit of our model (11) for the S&P index is virtually identical to that reported in (12) or (13) for the NYSE. For OLS and GLS, respectively, it is, for the period 1928-1980:

Figure 2. Schematic or box-plot of the distribution of deviations in annual percentage real dividend changes around their expected growth rate,<sup>1</sup> and percentage unexpected real price changes,<sup>1</sup> for the value-weighted NYSE companies over the period 1927-1980.



(Figure continued.)

Figure 2. The following is provided to aid in interpreting the (Cont'd.) above graph.



Definitions:

The upper hinge (UH). If the number of observations (NOB) divided by 4 equals an integer  $i$ , then  $UH = (X_{(3i)} + X_{(3i+1)})/2$ . But if  $NOB/4 = i$ , a non-integer, let  $j = [i]$ ; then  $UH = X_{(3j)}$ . Note that  $X_{(i)}$  and  $X_{(j)}$  are order statistics.

The lower hinge (LH). If  $NOB/4 = i$ , an integer; then  $LH = (X_{(i)} + X_{(i+1)})/2$ . But if  $NOB/4 = i$ , where  $i$  is a noninteger, let  $j = [i]$ ; then  $LH = X_{(j)}$ .

The midspread (MIDSP) is the distance  $UH - LH$ .

The upper side value is the largest data value less than  $UH + MIDSP$ . But if that value is less than UH, the upper side value equals UH.

The lower side value is defined similarly to the upper side value.

Outside points are data values between  $UH + (3/2) \cdot MIDSP$  and  $UH + MIDSP$  or between  $LH - (3/2) \cdot MIDSP$  and  $LH - MIDSP$ .

Detached points are data values greater than  $UH + (3/2) \cdot MIDSP$  or less than  $LH - (3/2) \cdot MIDSP$ .

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{-0.078}{(0.161)} + \frac{0.429}{(0.069)} \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] \quad (20)$$

$$-0.036 \log \frac{D(t)}{P(t-1)} + u(t+1) \\ (0.051)$$

$$\bar{R}^2 = 0.42 \quad DW = 1.80$$

and

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{-0.140}{(0.178)} + \frac{0.426}{(0.069)} \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] \quad (21)$$

$$-0.056 \log \frac{D(t)}{P(t-1)} + u'(t+1) \\ (0.061)$$

$$\bar{R}^2 = 0.41 \quad DW = 1.94$$

If the change in accounting earnings from year  $t$  to  $t+1$  is used to explain (not predict) the dividend change from year  $t$  to  $t+1$ , while  $(u - D/P)$  is used to account for the expected growth in accounting earnings, we obtain by OLS and GLS, respectively:

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{0.136}{(0.162)} + \frac{0.437}{(0.079)} \log\left[\frac{E(t+1)}{E(t)} + \frac{D(t)}{P(t)}\right] \quad (22)$$

$$+0.033 \log \frac{D(t)}{P(t-1)} + u(t+1) \\ (0.052)$$

$$\bar{R}^2 = 0.37 \quad DW = 1.80$$

and

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = \frac{0.146}{(0.158)} + \frac{0.518}{(0.081)} \log\left[\frac{E(t+1)}{E(t)} + \frac{D(t)}{P(t-1)}\right] \quad (23)$$

$$+0.039 \log \frac{D(t)}{P(t-1)} + u'(t+1)$$

(0.051)

$$\bar{R}^2 = 0.44 \quad DW = 2.06$$

In (22) and (23),  $E(t)$  refers to the aggregate accounting earnings for the S&P companies over year  $t$ . These results suggest that data on year  $t+1$  accounting earnings, which only become available the end-of-year  $t+1$ , explain no higher a percentage of aggregate dividend changes in year  $t+1$  than is explained by price data available at the beginning of year  $t+1$ . Moreover, the coefficient on the dividend-to-price ratio (i.e., the dividend-to-permanent-earnings ratio), has the wrong sign for regressivity, although that coefficient is imprecisely measured with only roughly fifty years of data. Except for the regressivity, the contemporaneous accounting earnings model has roughly the same fitted characteristics as the lagged stock price model.

If the accounting earnings change from period  $t-1$  to  $t$  is substituted for the period  $t$  to  $t+1$  change, the accounting model's explanatory power drops to about 10% (about 20% if a GLS error model is superimposed). An examination of the distributed leads and lags of dividend changes and accounting earnings changes confirms that the substantial portion of their association is, in fact, contemporaneous. This result is consistent with Fama and Babiak's (1968) finding that their "best" accounting earnings-based model of dividend changes includes contemporaneous and lagged earnings levels, since our accounting earnings change variable involves both contemporaneous and lagged accounting earnings levels.

Contemporaneous accounting earnings are roughly on a par with lagged stock price changes in explaining aggregate dividend changes, in part because these lagged stock price changes provide reasonably good forecasts of the subsequent earnings themselves. If the component of contemporaneous accounting earnings changes which could have been predicted from the lagged stock price change is removed, then the unpredictable component of contemporaneous earnings, which we denote by  $UE(t)$ , does add significantly (at the 95% level) to past price changes in explaining aggregate dividend movements. Using GLS over the 1929-1979 period:

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t-1)} = -0.069 + 0.426 \log\left[\frac{P(t) + D(t)}{P(t-1)}\right] \\ (0.166) \quad (0.069) \\ + 0.246 UE(t) - 0.032 \log\left[\frac{D(t)}{P(t-1)}\right] + u'(t+1) \quad (24) \\ (0.085) \quad (0.053)$$

$$\bar{R}^2 = 0.47 \quad DW = 1.88$$

The F statistic for inclusion of the earnings forecast error  $UE(t)$  is about 7.96, which is significant at the 5% level. The OLS estimates and GLS estimates of (24) are virtually identical, which is consistent with the earnings forecast error eliminating part of the serial dependence in the disturbances. The measured correlation between the variable  $\log[P(t+1)/P(t) + D(t+1)/P(t)]$  which, up to a constant, is the unexpected price change from the end of period  $t$  to the end of period  $t+1$ , and the (GLS) residual  $u(t+1)$  in (19) is about 10%. By including contemporaneous unanticipated earnings, this correlation between unexpected price changes and the (GLS) residual  $u(t+1)$  in (24) is reduced to about 3.3%. Thus, the already small amount of "information content" in aggregate dividends is further reduced by taking account of earnings changes. We pursue this no further because, as we noted earlier, the information content of dividends and earnings is probably not well defined for an aggregate of corporations over a coarse annual grid.



### 6.3 The Behavior of Gross or Net Dividends?

The previous model development and analysis has been focused on the behavior of aggregate gross dividends because, as Miller and Modigliani (1961) made clear, it is the determination of gross dividends which is at issue in the "dividend puzzle." We have, however, also fit our model to aggregate net dividends measured by aggregate gross dividends minus net new stock issues, and found the model's explanatory power and the estimates of its coefficients to be virtually unchanged. Given the cash flow identity that net cash flows equal net dividends, this finding might, for some, raise the question of whether our model really does describe the behavior of dividends, or whether it instead describes the behavior of net cash flows and investment policies.

However, for this standard measure of aggregate net dividends, there are two important sources of slippage in empirical applications of the cash flow identity which explain the similarity in behavior of gross and net dividends. Most obvious are cash inflows and non-investment outflows not accounted for in this measure of net dividends. These include changes in publicly issued corporate debt, privately placed corporate debt, bank loans, trade credit and other short term accruals, lease contracts, and legal and implied liabilities such as pensions and customer warranties.

New stock issues pale by comparison with these unaccounted-for sources and uses of funds. For example, a ten percent swing in the approximately \$120 billion of nonfinancial corporate net trade credit would alone amount to more than the average annual public offerings of common stocks of \$8 - 10 billion. Annual public offerings of corporate debt alone also typically exceed annual new equity issues. Thus, particularly in aggregate, it seems clear that dividends net of new equity issues cannot be reliably used to infer variations in net cash flows or investment policy.



Moreover, even if one were able to account for all the sources of leakage in the cash flow identity, the behavior of net dividends may still be similar to that reported in previous sections if there is enough flexibility in real investments to smooth out erratic movements in cash flow realizations. Casual inspection suggests that there exist many sources of such flexibility. For example, managers--as suppliers of dividends, and investor-consumers--as demanders, can accept, reject, or delay zero NPV real investments, choose among mutually exclusive investments whose NPV is about the same<sup>32</sup>, invest in noncorporate and overseas securities, etc. Again, the magnitudes involved speak for themselves. Annual corporate investment outlays are around \$300 billion, while aggregate corporate dividends total only \$60 billion. Thus a swing in net cash flows which would, with a fixed level of real investment, occasion (say) a twenty percent swing in net dividends, could be offset by a four percent change in real investment.

## 7. Conclusion

In this paper, we have shown that aggregate dividends exhibit a systematic time series behavior which can be well described by an error-correction-model in which aggregate real dividend changes are driven by the one-period lagged real changes in stock prices. Although the time series of aggregate dividends is considerably less volatile than the stock price series are, this model significantly outperforms the trend-autoregressive model where a distributed lag of past dividends together with a time trend is used to explain subsequent dividend changes.

The stock price model performs on a par with dividend models which use contemporaneous and lagged accounting earnings variables, as in previous studies of the Lintner (1956) model by Brittain (1966), Fama and Babiak (1966), and Watts (1973). However, because the stock price model uses only lagged prices, it can be used to forecast future dividend changes, whereas the accounting earnings model which employs contemporaneous earnings cannot be used to forecast. The version of the accounting earnings model which uses only lagged accounting earnings to forecast dividends significantly underperforms the stock price model.

As noted at the outset, a wide range of possible microtheories of dividend behavior could be consistent with the observed systematic behavior of aggregate dividends. However, our finding that aggregate dividends do exhibit systematic time-series behavior provides evidence that strictly firm-specific theories of dividends such as signalling, cannot by themselves explain the dividend puzzle.

FOOTNOTES

- \* We dedicate this paper to the scientific contributions and the memory of John V. Lintner, Jr.
- † We are grateful to J. Hausman, M. Miller, S. Myers, R. Ruback, and especially F. Black for helpful comments, and to the Institute for Quantitative Research in Finance for partial funding. The first author is also grateful to the Batterymarch Fellowship Program under which portions of the work here were completed.
- 1 See Brealey and Myers (1984, Ch. 16) for a more complete summary of the dividend controversy. The degree to which this controversy is unresolved is exemplified by its inclusion in Brealey and Myers (1984, p. 790) list of ten important unsolved problems in finance which "...seem ripe for productive research."
- 2 In fact, in their events study of the dividend behavior of split stocks, Fama, Fisher, Jensen, and Roll (1969) did adjust individual firm dividend changes for market-wide dividend changes before classifying the former as increases or decreases. FFJR did this to be internally consistent in associating security return residuals with dividend changes. Note that, as would be the case for our model at the micro-level, an added variable might well be required to account for dividend increases or decreases associated with a firm's "normal" progression through its "life cycle."
- 3 Without such an assumption, the Arrow-Debreu, Capital Asset Pricing Model, Arbitrage Pricing Model, and other spanning models used to establish equilibrium relations among securities and asset prices, would in general be empty since the number of states or systematic factors required would almost always be greater than or equal to the number of securities outstanding; see Merton (1982) discussion.
- 4 Easterbrook (1984) proposes that dividend payments might exist in part because they constitute one means of depleting internally generated funds, and thereby a means of forcing managers to go to the external market for investment funds (that market being assumed to be a low cost monitoring device). The degree to which a firm's dividends will be instrumental in "causing" the need for outside funds will, in this explanation, also depend upon where the firm is in its life cycle of cash flow generation.
- 5 We have considered whether there might instead be a "present-day" rational tax-based equilibrium for aggregate (gross) dividends analogous to Miller's for aggregate debt. There are two factors which seem to suggest otherwise. First, many institutions which are not taxed on dividends are not taxed on capital gains either, and hence, have no particular preference for dividends. Nor do broker-dealers who are taxed equally on both. Corporations might appear to prefer dividends to capital gains, but other instruments such as convertible preferreds make their demand for common stock dividends very fragile. Thus, both the aggregate demand and

supply curves for dividends may be very elastic in a tax-based model, and thus would not produce a sharp determinate equilibrium--it is in just such a context that our nonrational explanation for dividends makes most sense. Second, the supply-side adjustment mechanism for aggregate dividends would not be "as neat as" that in Miller's Debt and Taxes paper, where firms can observe both the tax rate on equity (i.e., the corporate tax rate in the simplest case), and the tax rate on the marginal dollar of debt (by comparing the before-tax cost of that marginal dollar of debt with municipal rates). Here, when considering supply-side adjustments, firms could only indirectly determine marginal investor dividend tax rates relative to capital gain tax rates by observing the effect of perturbations in their supply of dividends and capital gains on firm values, all else held constant.

- 6 The stylized facts distilled by Lintner from his interviews can be interpreted as a description of "average" or "systematic" dividend behavior. In this sense, these facts are macro rather than micro.
- 7 Although  $V(t)$  is the present value per share of the future cash flows available for distribution to the shares outstanding at time  $t$ , it does not follow that the dividend per share paid at time  $s$  must equal  $[\Pi(s) - I(s)]/N(s)$ . By the accounting identity,  $\Pi(s) - I(s) \equiv N(s)D(s) - [N(s+1) - N(s)]P(s)$  where  $[N(s+1) - N(s)]P(s)$  is the cash flow received from the issue of new shares of stock at time  $s$ , and therefore,  $D(s) = [\Pi(s) - I(s) + (N(s+1) - N(s))P(s)]/N(s)$ . If issues or purchases of shares are made at "fair" market prices, then such future transactions have a zero net present value, and therefore, have no effect on the current intrinsic value of the firm. If the firm has debt in its capital structure, then interest payments must be subtracted and net proceeds of new debt issues added to the cash flows of the firm. As with stock issues, if the debt is issued or retired at fair market prices, then such future debt transactions will also have no effect on the current intrinsic value of the firm. Although the additional future cash flows from new share and debt issues do not affect current permanent earnings, for a given value of permanent earnings, such financial transactions provide management with considerable flexibility to control the time path of dividends per share.
- 8 At least at the time of his survey thirty years ago, Lintner's evidence indicated that dividend policy was not viewed by management as simply a balancing item in the flow of funds account: "Dividends [rather than retained earnings and savings] represent the primary and active decision variable in most situations" (p. 197) "...In general, management's standards with respect to its current liquidity position appeared to be very much more flexible than its standards with respect to dividend policy, and this flexibility provided by the buffer between reasonably definite dividend requirements in line with established policy and especially rich current investment opportunities" (p. 105). Other statements could be interpreted as hints regarding the loss function underlying dividend rules, but they are not very specific, e.g., by "stabilizing" dividends, managers can "minimize adverse shareholder reactions," and "management can live more comfortably with its unavoidable uncertainties regarding future developments" (p. 100).



- 9 It may be verified that only trivial modifications in the dynamical equation developed below are required to explicitly account for any steady-state variation in the ratio. For example, if it is hypothesized that the equilibrium ratio depends upon a vector of stationary stochastic variables  $\underline{Z}(t)$ , (5) may be replaced by  $\log[D(t)/E^m(t)] = \beta_0 + \beta' \underline{Z}(t)$ , the only effect being a change in the regressivity term in (5) below.
- 10 The interpretation of the parameter  $\gamma$  in (5) suffers from the absence of a more precise underlying structural model which leads to (3). To see this, suppose for simplicity that  $\Delta \log[E^m(t)] = m(t-1)$ ,  $\eta(t) = 0$ , in (5), but that the current dividend yield is out-of-equilibrium. Salmon (1982, p. 622) shows that a "proportional, integral, derivative" (PID) control rule:

$$\begin{aligned} \Delta \log[D(t)] = & \beta - k_i \{ \log[D(t)] - \log[E^m(t-1)] \} - k_p \Delta \{ \log[D(t)] \\ & - \log[E^m(t-1)] \} - k_d \Delta^2 \{ \log[D(t)] - \log[E^m(t-1)] \} \end{aligned} \quad (i)$$

leads to the following error correction term:

$$\begin{aligned} \Delta \log[D(t)] = & \{ (k_p + k_i + k_d) - (k_p + 2k_d)L - k_d^2 L^2 \} \{ \log[D(t)] - \log[E^m(t-1)] \} \\ & \equiv \{ \beta - A(L) \{ \log[D(t)] - \log[E^m(t-1)] \} \} \end{aligned} \quad (ii)$$

In the absence of a more explicit model, one can only speculate about the need to include a term like  $\Delta^2 \{ \log[D(t)] - \log[E^m(t-1)] \}$  in (5). For example, in Salmon's general discussion, he argues that an error correction rule might be appropriate when a decision-maker faces an uncertain environment in which the control problem is itself changing over time. Terms like  $\Delta^2 \{ \log[D(t)] - \log[E^m(t-1)] \}$  may pick up such changes, especially if the type of model changes that occur result in a non-time-additive "control problem" to be solved for the short-run dynamics of dividends. Insofar as changes in the long-run equilibrium dividend-price ratio are concerned, many can be accommodated in the formulations in the text. Given these formulations, if uncertainties about the "short-run" dividend control problem can not be completely described by a linear-quadratic problem, a non-linear model will generally be needed in place of (5)--the certainty equivalence principle will no longer hold and  $\eta(t)$  will not be an additive error.

- 11 Strictly, feasibility only requires the weaker "less than or equal to." If, however, dividends include all distributions to stockholders and if managers do not throw cash away, then strict equality is required. In contrast to the actual dividend payments made, the term "dividend policy" refers to the contingent schedule or plan for future dividend payments. A dividend policy is, thus, much like the state-contingent functions for optimal control variables which are derived from the solution of a stochastic dynamic programming problem.

- 12 Indeed, even in the restrictive context of our simple behavioral model, any values for  $\lambda$  and  $\gamma$  in (10) such that  $0 \leq \lambda \leq 1$  and  $\gamma > 0$  are more than sufficient to ensure satisfaction of the rationality constraint.
  
- 13 Even the strongest supporters of the view that "dividend policy matters" would agree that the only effect of a change in dividend policy on investment policy is through its effect on the firm's cost of capital,  $\alpha$ . Although a change in dividend policy may "signal" a change in investment policy, one could hardly argue that such a dividend policy change "caused" the subsequent change in investment policy that it signalled.
  
- 14 See, for example, Hall (1978). We note further that if consumer behavior is to smooth the time path of changes in consumption, then the dynamics for a change in next period's consumption in response to an unanticipated change in this period's wealth may well be described by a partial adjustment process analogous to our equation (5).
  
- 15 Note that the degree of market efficiency posited here is much weaker than would be implied by assuming that the market information set contains all the relevant information contained in managers' aggregated information sets. Under our assumption, a manager may have information relevant to the estimation of his or her firm's intrinsic value that is not available to the market. If, as would seem reasonable, such nonpublic information is firm-specific, then differences between the market's and the manager's assessment of the individual firm's intrinsic value that arise from this source are likely to (statistically) disappear when these individual assessments are averaged over all firms. It is, of course, possible that the market's information set is richer than the individual manager's, even with respect to estimates of his or her own firm's intrinsic value. However, rationally-behaving managers would presumably take this possibility into account when making their dividend decisions.
  
- 16 As discussed in footnote 7, because of transactions by the firm in its own liabilities, it is not the case that  $[\Pi(s) - I(s)]/N(s) \equiv D(s)$  in (8). Even without such transactions, managers can still implement virtually any change in dividends per share by the purchase or sale of financial assets held by the firm or by marginal changes in the amount of investment in any other "zero net present value" asset (e.g., inventories). While these latter transactions will change the time pattern of  $\Pi(s) - I(s)$ , they will not affect the present value of these future cash flows, and therefore, will not affect the current level of permanent earnings.
  
- 17 If the stock market is moved partly by irrational waves of optimism and pessimism, (9) will be replaced by:  

$$\log[P_c(t)/P_c(t-1)] = \psi \log[E^m(t)/E^m(t-1)]$$

where  $\psi > 1$ , and the coefficient  $\lambda$  of the stock price change term in (10) will be replaced by  $\lambda/\psi$ . Thus,  $a_1 < 1$  in the reduced

form (11) would be consistent with immediate and full adjustment of dividends to permanent earnings by managers (i.e.,  $\lambda = 1$ ) and  $\psi > 1$  (the market is irrational). Note that if managers set dividends in accord with rationally assessed permanent earnings, and irrational stock prices eventually revert to their rational levels, the dividend-price ratio would be a predictor of subsequent cum-dividend price changes. However, the dividend-price ratio will also predict price changes if expected permanent earnings changes vary over time in a serially dependent way. We discuss the empirical relation between dividend yield and subsequent stock returns in Section 6.1.

- 18 We have also fitted (17) using quarterly data. Although it might at first appear that the use of quarterly rather than annual data would quadruple the number of observations available, there are good reasons for doubting this. There is a distinct yearly (and half-yearly) seasonal in real quarterly dividends. If, as this suggests, managers wait until the fourth (fiscal) quarter to "take a look at the year's performance" before deciding to raise or lower that year's dividend relative to the previous year's, then the last quarter's dividend contains effectively the same information as the annual dividend. Further, any "seasonal adjustment" of quarterly dividends not only runs the risk of smoothing away the very innovations in dividends in which we are interested, but also is doubly hazardous when autocorrelated disturbances or lagged dependent variables might be present as in (11). We therefore report only the results for annual data.
- 19 We note that our data set is different from those used by Shiller. However, using our data set, we were able to obtain essentially the same empirical findings as reported by Shiller for his data sets. We, therefore, expect that the results reported here for our model will also obtain if it were fit to his data sets.
- 20 Any of the above-mentioned ways in which the adjustment lags could arise are also potential explanations for disturbance autocorrelation, because in dynamic regression models, lag structure and disturbance autocorrelation can act as proxies for each other.
- 21 Our  $R^2$  is below the 85% figure given in the original Lintner (1956) study. However, Lintner (1956), who stressed that his results were preliminary, pooled his time series observations from 1918-1941 for his 28 companies to estimate his model. Fama and Babiak (1968), who reestimated Lintner's model separately for each firm over the years 1946-1964, report (e.g., in their Table 2) average  $R^2$  figures of (roughly) 40%-45% which are comparable with ours.
- 22 The point estimate of the dividend yield coefficient is sensitive to how the autocorrelation is taken into account. The half-life calculation in the text uses the GLS rather than the OLS estimate because the positive autocorrelation in the OLS residuals reduce the absolute magnitude of the (negative) dividend yield coefficient. Dynamic regression models with autocorrelated disturbances cannot be easily distinguished from ones with lagged dependent variables, and the dividend yield will be correlated with the lagged dependent variable if, as both our model and the Shiller



model posit, the dividend-to-price ratio is autoregressive. There is evidence that the residual autocorrelation estimate is somewhat sensitive to an outlier in 1951, but this outlier apparently has no effect on the estimates of the coefficients in our model. We therefore omit a more detailed analysis of the disturbance autocorrelation.

- 23 In their classic events study of stock splits and the cash dividend changes which often accompany them, Fama, Fisher, Jensen, and Roll (1969) report a result similar to our's: stock splits, and the increased dividends which typically accompany them, were on average preceded by abnormal price increases. In their study, some of the average "run-up" in prices took place earlier than twelve months before the stock split event, but FFJR deliberately select the individual companies which, ex post, split their stocks. The early small run-up in prices could easily wash out in our aggregate data, and in any case, the FFJR study does not provide information on changes in cash dividends other than those associated with the stock split.
- 24 In previous estimations, we used a "one-step" GLS procedure. Because Maddala (1971) has shown that "iteration pays" in GLS estimation when a lagged dependent variable is present, we use the iterative approach for the equations in this section.
- 25 Comparing the OLS and GLS  $R^2$ s provides only a heuristic measure of the incremental explanatory power afforded by the GLS regression, because the OLS  $R^2$  is not a proper benchmark in light of autocorrelation in the residuals. The  $R^2$  for the GLS regression, which, in this case and all others in the paper, we compute as (geometrically) the square of the cosine of the angle between the (centered) dependent variable and the (centered) fitted dependent variable, is also well known not to be uniquely defined for GLS and nonlinear regression models. However, we believe our statements in this paper concerning model fit are not sensitive to our  $R^2$  measure, especially since the OLS and GLS fits of our model are essentially the same. Further, it is hard to think of a more "natural" way of generally measuring the tightness between the fitted and actual dependent variables.
- 26 As was the case for the OLS and GLS fits (12) and (13) of our model, the GLS fit of (16) does not significantly improve upon its OLS fit. The F statistic for the autoregressive correction in (16) is 2.85, which is not significant at the conventional 5 percent level. Thus, the addition of log price change to (15) substantially increases its explanatory power and improves the autoregressive model's specification (14) by eliminating the requirement that it be supplemented by more structure on the stochastic process for the "unknown" variables before it is a proper regression equation.
- 27 This regression is almost the equivalent of the Granger-Sims causality test referred to in the causality discussion in Section 3.
- 28 It is all the more surprising because the model does not use contemporaneous price changes. Because all the variables in our model are lagged, equation (18) is a "true" forecast equation in the sense that at time  $t$ , it provides an unconditional forecast  $D(t + 1)$ . The

relatively high  $\bar{R}^2$  suggests that aggregate dividends may be forecasted rather successfully.

- 29 There is, of course, ample precedent in the economics literature for assuming that relative values such as the dividend-to-price and earnings-to-price ratios have steady-state or stationary distributions. As exemplified by our model, the existence of steady-state distributions for such relative values surely does not justify the assumption of stationarity distributions for the levels (or absolute values) of dividends, earnings, or prices. Further, unless investors are risk-neutral and the riskless rate of interest is constant, the assumption of a constant expected return on the market, made in Shiller (1981a,b) and in much of the finance literature, is inconsistent with the trend-autoregressive process for dividends and stock prices. For proofs and further discussion, see Myers and Turnbull (1977) and Fama (1977).
- 30 Study of our sample using blunt interocular analysis suggests that there are a few influential "outliers" in the annual data which cause the correlation, and the correlation disappears all together with monthly data.
- 31 We also used  $\alpha(1-D/E)$  with  $\alpha = 0.048$  as a measure of this expected growth. However, the results obtained are no different to those reported.
- 32 This set includes the projects which can be delayed a year with little or no penalty, where the penalty must be assessed net of timing and abandonment options which are gained by the delay.

# APPENDIX A

Virtually any sensible economic model will possess the property that variables such as the real interest rate and the aggregate dividend-yield ratio tend to a steady-state. This requirement is of course satisfied for dividend yields if the levels of (detrended) dividends and prices are themselves stationary (e.g., Shiller (1981a,b)). Here we show that it can also be met in our dividend model when price levels are not stationary.

We assume, as did Shiller, that the expected real rate of return on a share of the market portfolio, denoted by  $\alpha$ , is constant over time. However, we conjecture that our conclusions concerning the stationarity of the dividend-price ratio will not be substantively affected if the expected real rate is not constant, so long as movements in that rate follow a steady-state distribution.

If  $P(t)$  denotes the price of a share in the market portfolio at time  $t$  (again, in real terms), then by the well-known valuation equation,  $P(t)$  can be written as:

$$P(t) = \epsilon_t^M \left\{ \int_t^\infty e^{-\alpha(s-t)} D(s) ds \right\} \quad (A1)$$

where  $D(s)$  denotes aggregate real dividends per share, and  $\epsilon_t^M$  is the expectation operator over the probability distribution for future dividends generated by the information set available to investors in the market as of time  $t$ .

From (A1), we derive the dynamic equation for the evolution of stock price which can be separated into two components: the expected change and the unanticipated change. The expected change in the price between time  $t$  and  $t + h$  can be written as:

$$\epsilon_t^M \{P(t+h) - P(t)\} = [k(h)P(t) - \bar{D}(t;h)]h \quad (A2)$$

where  $k(h) \equiv (e^{\alpha h} - 1)/h$  is the expected rate of return on the stock

over the interval  $h$  and  $\bar{D}(t;h) \equiv e^{\alpha h} \epsilon_t^M \left\{ \int_t^{t+h} \exp[-\alpha(s-t)] D(s) ds \right\} / h$

is the expected end-of-period value of cumulative dividends paid during the period expressed as a flow per unit of time. From (A1), the unanticipated component of the change in price can be written as:

$$P(t+h) - P(t) - \epsilon_t^M \{P(t+h) - P(t)\} \quad (A3)$$

$$= \int_{t+h}^{\infty} \exp[-\alpha(s-t-h)] \{ \epsilon_{t+h}^M [D(s)] - \epsilon_t^M [D(s)] \} ds$$

which reflects the revision between  $t$  and  $t+h$  of the market's expectation of future dividends. Note that from the perspective of any date equal to or earlier than  $t$ , the expected value of the right-hand side of (A3) is always zero. Hence, it follows, by construction, that the sequence of unanticipated price changes described by (A3) form a Martingale.

Any stock price process posited for (A2) and (A3) must satisfy the restriction that share owners enjoy limited liability (i.e.,  $P(t) \geq 0$ ). Moreover, to avoid arbitrage,  $P(t) = 0$  if and only if  $D(s) \equiv 0$  for all  $s > t$ . We can, therefore, rule out a priori the arithmetic random walk model for stock price change. Because  $k$  is constant and  $\bar{D}(t;h)$  is nonnegative, the unanticipated change in price in (A4) cannot be described by a Gaussian distribution with a constant variance. Indeed, the variance of the unanticipated change in price must depend on the level of price in a systematic way. Therefore, any attempt to use standard regression analysis to fit the arithmetic change in price will not only produce inefficient estimates

because of heteroscedasticity, but will also produce biased estimates of the coefficients for those explanatory variables which are correlated with the level of  $P(t)$ .

Given that prices must be positive, it is more useful to express the price dynamics in terms of percentage change. From (A2) and (A3), we have that

$$\frac{P(t+h) - P(t)}{P(t)} = q(t;h)h + w(t;h) \quad (A4)$$

where  $q(t;h) \equiv [k(h) - \bar{D}(t;h)/P(t)]$  is the expected growth rate in price between  $t$  and  $t+h$  and  $w(t;h)$  is the unanticipated percentage change in price which is equal to the right-hand side of (A3) divided by  $P(t)$ .

To locate (A4) within the financial economics literature, we note that the standard continuous-time version of (A4) is a stochastic differential equation of the  $\hat{\text{Itô}}$  type which we write as:

$$\frac{dP}{P} = q(t)dt + \sigma dz \quad (A5)$$

where  $q(t) \equiv \alpha - D(t)/P(t)$ ;  $\sigma$  is the instantaneous standard deviation of the rate of return on the stock; and  $dz$  is a standard Wiener process. The prototype version of (A5) assumes both  $\alpha$  and  $\sigma$  are constants which implies that total returns on the stock (i.e., dividends plus price appreciation) follow a geometric Brownian motion.

For analytical simplicity, we work with (A5) rather than (A4). That is, we assume that  $w(t;h)$  in (A4) can be expressed as:

$$w(t;h) = \frac{1}{P(t)} \int_t^{t+h} \sigma P(s) dz \quad (A6)$$

where  $\sigma$  is assumed to be a constant.

While we assume (A6) for analytical convenience, it does not create any



material "model bias" with respect to the conclusion below or any substantive issues of the paper. As we have already noted, any posited process for stock price must satisfy the Martingale property for  $w(t;h)$ , and therefore, a nonanticipating process like a Brownian motion is almost a necessity for consistency. Although the assumption of (A5) with a constant  $\sigma$  implies that cumulative total returns over a given interval are lognormally distributed with a variance rate which grows with the length of the interval, the stock price itself need not be lognormally distributed because of the "dividend-drag" term in (A5). Indeed, depending upon the properties of the dividend process, (A5) is entirely consistent with stock price having a steady-state or long-run stationary distribution around some fixed time trend.

From equation (A5), we have that the price dynamics can be written as:

$$\frac{dP}{P} = (\alpha - D/P)dt + \sigma dz \quad . \quad (A7)$$

By Ito's lemma, it follows from (A7) that:

$$d(\log P) = (\alpha - \frac{1}{2} \sigma^2 - D/P)dt + \sigma dz \quad . \quad (A8)$$

As shown, for example, in Merton (1975, Appendix B), if  $\log P$  has a nondegenerate steady-state distribution, then there must exist numbers  $\bar{P} < \infty$  and  $\underline{P} > 0$  such that for (almost) all  $P \geq \bar{P}$ ,  $(\alpha - \frac{1}{2} \sigma^2 - D/P) < 0$  and for (almost) all  $P < \underline{P}$ ,  $(\alpha - \frac{1}{2} \sigma^2 - D/P) > 0$ .<sup>\*</sup> By hypothesis,  $\alpha$  and  $\sigma$  are constants. Hence, the full burden of these requirements falls upon the dividend-to-price ratio viewed as a function of price. As a practical matter, these requirements on  $D/P$  are likely to be met only if this ratio is a systematically increasing function of the price.



Empirically, the dividend-to-price ratio is not such a function, and therefore, the price will not have a steady-state distribution. Moreover, the same conclusion applies as well to "detrended" price. That is, let  $p(t) \equiv P(t)\exp[-gt]$  and  $d(t) \equiv D(t)\exp[-gt]$  where  $g$  is some appropriately chosen deterministic trend rate. From (A8), it follows that:

$$d(\log[p(t)]) = (\alpha - \frac{1}{2}\sigma^2 - g - d/p)dt + \sigma dz \quad (A9) \quad .$$

Because  $d(t)/p(t) = D(t)/P(t)$ , the conditions on  $D/P$  for detrended price to have a steady-state distribution are the same as for the nondetrended price except the constant " $\alpha - \sigma^2/2$ " is replaced by another constant " $\alpha - \sigma^2/2 - g$ ."

If, for analytical convenience, we neglect the empirically-established lag in the dividend response to an unanticipated change in price, then the continuous-time version of our reduced-form equation (5) can be written as:

$$\frac{dD}{D} = (\alpha - \frac{D}{P} + \gamma[\beta - \log(\frac{D}{P})])dt + \lambda\sigma dz \quad (A10)$$

and by Ito's lemma, we have from (A10) that:

$$d(\log D) = [\alpha - \lambda^2\sigma^2/2 - \frac{D}{P} + \gamma[\beta - \log(\frac{D}{P})])dt + \lambda\sigma dz \quad (A11) \quad .$$

By subtracting (A8) from (A11) and rearranging terms, we can write the dynamics for the dividend-to-price-ratio as:

$$d[\log(D/P)] = \gamma[\mu - \log(D/P)]dt - (1 - \lambda)\sigma dz \quad (A12)$$

where  $\mu \equiv \beta + (1 - \lambda^2)\sigma^2/2\gamma$ . By inspection of (A12),

$\log[D/P]$  follows a classical Ornstein-Uhlenbeck process which is known to have a steady-state normal distribution with mean  $\mu$  and variance  $(1 - \lambda)^2\sigma^2/2\gamma$ . It follows from (A12) that the dividend-to-price

ratio has an asymptotic stationary lognormal distribution.

Although the dividend-to-price ratio has a steady-state distribution, the level of stock prices does not. It therefore follows that the level of dividends cannot have an asymptotic stationary distribution. Similarly, from the analysis of detrended price in (A9), the level of detrended dividends,  $d(t)$ , does not have a stationary distribution.

- \* See Merton (1975, Appendix B). The proof depends on the behavior of the dividend-to-price ratio in the asymptotic regions where  $P$  is either very large or very small. As a strict mathematical condition, it is not necessary that  $D/P$  be monotonically increasing in these regions for a nondegenerate steady-state distribution to exist. However, while "local" violations of monotonicity are permitted, the overall "trend" of the function in those regions must be positive. We see no theoretical or empirical foundation for assuming that  $D/P$  behaves in this way. It is possible that  $D/P$  is, in fact, an increasing function of  $P$  in these regions, but that the values of  $P$  in the 1926-1981 sample period were not observations from these regions. If so, then, of course, price will have a asymptotic stationary distribution. However, in that case, it is readily apparent that it may take hundreds or perhaps thousands of years to observe it.

# APPENDIX B

The long-run steady state (4) in the text can be imposed on the short-run dynamics (3) by the further coefficient restrictions:

$$\lambda - \theta_1 = 1 - \phi_1 \equiv \gamma \quad (\text{i.e., } \lambda - \theta_1 = 1 - \phi_1) \quad , \quad (\text{B.1})$$

and by setting the constant in the deterministic function  $a(t)$  in (3) equal to  $\gamma\beta$ . Denoting the function  $a(t)$  sans the constant by  $a'(t)$  and incorporating (B.1), we can rewrite (3) as follows:

$$\begin{aligned} \log[D(t+1)] - \log[D(t)] &= \gamma\beta + a'(t) + \lambda\{\log[E^m(t)] - \log[E^m(t-1)]\} \\ &\quad - \gamma\{\log[D(t)] - \log[E^m(t-1)]\} + \eta(t+1) \end{aligned} \quad (\text{B.2})$$

or, by rearrangement:

$$\begin{aligned} \log[D(t+1)] - \log[D(t)] &= a'(t) + \lambda\{\log[E^m(t)] - \log[E^m(t-1)]\} \\ &\quad + \gamma\{\beta - (\log[D(t)] - \log[E^m(t-1)])\} + \eta(t+1) . \end{aligned} \quad (\text{B.3})$$

To derive (5) in the text, define the expected logarithmic change in permanent earnings,  $\{\log[E^m(t)] - \log[E^m(t-1)]\}$ , as  $m(t-1)$ . Then (B.3) can be rewritten as:

$$\begin{aligned} \log[D(t+1)] - \log[D(t)] &= a'(t) + \lambda m(t-1) + \lambda\{\log[E^m(t)] - \log[E^m(t-1)] - m(t-1)\} \\ &\quad + \gamma\{\beta - (\log[D(t)] - \log[E^m(t-1)])\} + \eta(t+1) . \end{aligned} \quad (\text{B.4})$$

Substituting  $g(t) \equiv a'(t) + \lambda m(t)$  gives (5) in the text.

REFERENCES

- Aharony, J., and I. Swary, 1980, "Quarterly Dividend and Earnings Announcements and Stockholders' Returns: An Empirical Analysis," The Journal of Finance 35(1), 1-11.
- Asquith, P., and D.W. Mullins, Jr., 1983, "The Impact of Initiating Dividend Payments on Shareholder's Wealth," Journal of Business, 56(1), 77-95.
- Bhattacharya, S., 1979, "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," Bell Journal of Economics, Spring, 259-270.
- Black, F., 1976, "The Dividend Puzzle," in S.C. Myers, ed., Modern Developments in Financial Management, Hinsdale, IL, Dryden Press.
- Black, F., and M.S. Scholes, 1974, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," Journal of Financial Economics, 1 (May), 1-22.
- Brealey, R.A., and S.C. Myers, 1984, Principles of Corporate Finance, 2nd ed., New York, McGraw-Hill.
- Brittain, J.A., 1966, Corporate Dividend Policy, Washington, DC, The Brookings Institution.
- Darling, P.G., 1955, "A Surrogate Measure of Business Confidence and Its Relation to Stock Prices," Journal of Finance, 10 (December), 442-458.
- Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo, 1978, "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom," The Economic Journal, 88, December, 661-692.
- Dickey, D.A., and W.A. Fuller, 1979, "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," Journal of the American Statistical Association, 74(366) (June), 427-431.
- Dickey, D.A., and W.A. Fuller, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," Econometrica, 49(4) (July), 1057-1072.
- Eades, K.M., P.J. Hess, and E.H. Kim, 1984, "On Interpreting Security Returns During the Ex-Dividend Period," Journal of Financial Economics, 13(1), 3-34.
- Easterbrook, F.H., 1984, "Two Agency Cost Explanations of Dividends," The American Economic Review 74(4), 650-659.

- Ericsson, N.R., and D.F. Hendry, 1984, "Assertion Without Empirical Basis: An Econometric Appraisal of Monetary Trends in ... The United Kingdom by Milton Friedman and Anna J. Schwartz," Paper presented to the NBER Research Program on Economic Fluctuations, Macro Conference, July 12-13, Cambridge, MA.
- Fama, E.F., 1977, "Risk-Adjusted Discount Rates and Capital Budgeting Under Uncertainty," Journal of Financial Economics, 5, 3-24.
- Fama, E.F., and H. Babiak, 1968, "Dividend Policy: An Empirical Analysis," Journal of the American Statistical Association, 63 (December), 1132-1161.
- Fama, E.F., L. Fisher, M. Jensen, and R. Roll, 1969, "The Adjustment of Stock Prices to New Information," International Economic Review, 10 (February), 1-21.
- Gonedes, N.J., 1978, "Corporate Signalling, External Accounting, and Capital Market Equilibrium: Evidence on Dividends, Income, and Extraordinary Items," Journal of Accounting Research \_\_ ( ), 16-79.
- Gordon, M.J., 1959, "Dividends, Earnings and Stock Prices," Review of Economics and Statistics, 41 (May), 99-105.
- Gordon, M.J., 1962, "The Savings, Investment and Valuation of a Corporation," Review of Economics and Statistics, (February), 37-51.
- Hall, R.E., 1978, "Stochastic Implications of the Life Cycle--Permanent Income Hypothesis: Theory and Evidence," Journal of Political Economy, 86(6), 971-987.
- Hausman, J.A., 1978, "Specification Tests in Econometrics," Econometrica, 46(6), 1251-1271.
- Hess, P.J., 1983, "Tests for Tax-Effects in the Pricing of Financial Assets," Journal of Business, Forthcoming.
- John, K., and J. Williams, 1984, "Dividends, Dilution, and Taxes: A Signalling Equilibrium," Unpublished Paper, Graduate School of Business, New York University, 100 Trinity Place, New York, NY 10006.
- Kleidon, A.W., 1983, "Variance Bounds Tests and Stock Price Valuation Models," Unpublished Working Paper, Graduate School of Business, Stanford University, (January).
- Lintner, J., 1956, "Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes," The American Economic Review, XLVI(2) (May), 97-113.
- Lintner, J., 1962, "Dividends, Earnings, Leverage, Stock Prices and The Supply of Capital to Corporations," Review of Economics and Statistics, XLIV(3) (August), 243-269.



- Litzenberger, R.H., and K. Ramaswamy, 1979, "The Effect of Personal Taxes and Dividends on Capital Asset Prices," Journal of Financial Economics, 7, 163-195.
- Maddala, G.S., 1971, "Generalized Least Squares With An Estimated Variance Covariance Matrix," Econometrica, 39(1), (January), 23-33.
- Malinvaud, E., 1966, Statistical Methods of Econometrics, Amsterdam: North-Holland Publishing Company.
- Marsh, T.A., and R.C. Merton, 1984, "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," Sloan School of Management Working Paper #1584-84, August.
- Marsh, T.A., and R.C. Merton, 1985, "Corporate Dividend Dynamics at the Firm Level," Unpublished Paper, Sloan School of Management, MIT, 50 Memorial Drive, Cambridge, MA 02139, April.
- Merton, R.C., 1975, "An Asymptotic Theory of Growth Under Uncertainty," The Review of Economic Studies, 62(3), July, 375-393.
- Merton, R.C., 1982, "On the Mathematics and Economic Assumptions of Continuous Time Models," W.F. Sharpe, Ed., Financial Economics: Essays in Honor of Paul Cootner, North-Holland, Amsterdam.
- Miller, M.H., 1977, "Debt and Taxes," The Journal of Finance, 32(2), May, 261-175.
- Miller, M.H., and F. Modigliani, 1961, "Dividend Policy, Growth and the Valuation of Shares," Journal of Business, October, 411-433.
- Miller, M.H., and M.S. Scholes, 1978, "Dividends and Taxes," Journal of Financial Economics 6, 333-364.
- Miller, M.H., and M.S. Scholes, 1982, "Dividends and Taxes: Some Empirical Evidence," Journal of Political Economy, 90, 1118-1142.
- Miller, M.H., and K. Rock, 1982, "Dividend Policy Under Asymmetric Information," Unpublished Paper, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637, November. Forthcoming: Journal of Finance.
- Myers, S.C., and S.M. Turnbull, 1977, "Capital Budgeting and The Capital Asset Pricing Model: Good News and Bad News," Journal of Finance, 32(2), 321-333.
- Nickell, S., 1980, "Error Correction, Partial Adjustment and All That: An Expository Note," C.L.E. Working Paper No. 255, London School of Economics, May.
- Petit, R.R., 1972, "Dividend Announcements, Security Performance, and Capital Market Efficiency," Journal of Finance, December, 993-1007.



- Poterba, J.M., and L.H. Summers, 1984, "The Economic Effects of Dividend Taxation," National Bureau of Economic Research Working Paper #1353, May.
- Plosser, C.I., G.W. Schwert, and H. White, 1982, "Differencing as a Text of Specification," International Economic Review, 23(3), October, 535-552.
- Salmon, M., 1982, "Error Correction Mechanisms," The Economic Journal, 92, September, 615-629.
- Sargan, J.D., 1964, "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in P.E. Hart, G. Mills and J.K. Whitaker, eds., Econometric Analysis for National Economic Planning, London, Butterworths, 25-63.
- Shefrin, H.M., and M. Statman, 1984, "Explaining Investor Preference for Cash Dividends," Journal of Financial Economics, 13(2), 253-282.
- Shiller, R.J., 1981a, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" The American Economic Review, 71(3), June, 421-436.
- Shiller, R.J., 1981b, "The Use of Volatility Measures in Assessing Market Efficiency," The Journal of Finance, XXXVI(2), May, 291-311.
- Shiller, R.J., 1983, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?: Reply," The American Economic Review, 73(1), 236-237.
- Theil, H., 1965, Linear Aggregations of Economic Relations, Amsterdam, North-Holland.
- Telser, L.G., 1967, "Discrete Samples and Moving Sums in Stationary Stochastic Processes," Journal of the American Statistical Association, 62(318), June, 484-499.
- Tukey, J.W., 1970, Explanatory Data Analysis, Addison-Wesley, Reading, MA.
- Watts, R., 1973, "The Information Content of Dividends," The Journal of Business, April, 191-211.
- White, N., 1980, "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," Econometrica, 48(4), May, 817-838.
- Zellner, A., 1979, "Causality and Econometrics," in K. Brunner and A.H. Meltzer, eds., Three Aspects of Policy and Policy Making, Carnegie-Rochester Conference Series, Volume 10, Amsterdam, North-Holland Publishing Co.



3 9080 003 060 677





BASEMENT	Date Due	
MAR 13 1990		Lib-26-67

Barcode is on back cover.



